12-1-2016

The Competing Risks of Prepayment and Default on the Single-Family Mortgage Market

Ran An
Syracuse University

Abstract

This study contains three chapters. Since the subprime crisis, it has become increasingly important to understand the competing risks of prepayment and default on the single-family mortgage market. This research studies the economic factors that affect the competing risks of prepayment and default in locations where the aggregate of the prepayment risk and the default risk are simultaneously high.

Chapter 1 outlines the analysis based on thirty-year, fixed-rate, single-family mortgages in five Metropolitan Statistical Areas (MSAs): Phoenix, Miami, Tampa, Detroit and Las Vegas. These MSAs are chosen from a sample Single-family Loan-Level Dataset constructed by Freddie Mac based on high simultaneous prepayment and default rates. The results are estimated by a discrete time competing risks model based on restricted multinomial logit. Two different combinations of dependent variables are used to make the analysis more comprehensive. The first combination is prepayment and default and the second is prepayment and 90-days-delinquency.

The indicator for a prepayment penalty and the value of the call option are used to evaluate the prepayment risk and the best three combinations (using the Bayesian information criterion) of explanatory variables – the value of negative equity, a negative equity dummy, and original loan-to-value ratio – together with the unemployment rate are used to evaluate the default risk. Because of the ambiguous effect of the credit score on the prepayment decision discussed in the previous literature, two estimations to explain the effect of the credit score on the prepayment and default are run. One estimation considers how the credit score affects both termination risks and the other considers how the credit score affects only default. The effect of the debt-to-income ratio on the prepayment and default is also tested.

By controlling other mortgage characteristics, such as loan size and loan age, the brief summary of results shows us the prepayment penalty has a negative effect on the prepayment decision. However, its effect in Detroit is insignificantly positive. The value of the call option has a significantly positive impact on the prepayment in each MSA and its effect is the strongest among all of the explanatory variables. The unemployment rate has a positive effect on default/90-days-delinquency. Moreover, negative equity, the negative equity dummy and original loan-to-value are all positively related to the default/90-days-delinquency decision. As expected, in most MSAs, the credit score has a strongly positive effect on the prepayment; comparatively, it has a strong negative impact on the default. When using the credit score only in the default risk and not in the prepayment risk, the effect of the credit score is still significantly negative, but the coefficients decrease slightly. Moreover, debt-to-income does not appear to affect prepayment; however, it has a positive relationship with the default/90-days-delinquency in most MSAs. A Wald test is constructed to test the equality of the coefficients among five MSAs and the results of this test support the argument that prepayment and default risks are heavily influenced by local characteristics.

In earlier studies, the proportional hazards model becomes a popular method to analyze the single (prepayment or default) risk of mortgages. This model has been developed into a competing risks model and been widely used to analyze the prepayment and default risks simultaneously in later studies. However, the process of constructing the competing risks model is ambiguous. In Chapter 2, this study clearly presents the calculation process of this model based on the proportional hazards model using Sueyoshi’s method and then implements the model to analyze the termination risks of single-family mortgages in Phoenix.

Ding, Tian, Yu and Guo (2012) construct a new model based on a class of transformation survival models to analyze the risk of bankruptcy and they argue that that the proportional hazards model is not the best model to analyze this risk. Therefore, a question is raised by this argument: whether the proportional hazards model is the best model to analyze the default/prepayment risk of single-family mortgages? A new competing risks model based on a class of discrete transformation survival models is constructed in Chapter 2 and it is used to analyze the termination risks of the single-family mortgages in Phoenix. The model is controlled by the transformation parameters c_p (for the prepayment risk) and c_d (for the default risk). When $c_p = 0$ and $c_d = 0$, it is the competing risks model based on proportional hazards, and when $c_p > 0$and $c_d > 0$, its framework is changed according to the value of c_p and c_d.

The results show that the proportional hazards framework is the best model to estimate the prepayment risk, but it is not the best model to estimate the default/90-days-delinquency risk. The results of both models support the important arguments made in Chapter 1. Comparing the coefficients estimated by three competing risks models, the coefficients estimated by the model based on the Sueyoshi proportional hazards are insignificantly distinguishable from those estimated by the model based on the multinomial logit. Moreover, the coefficients estimated by the model based on a class of transformation survival models are significantly different from those estimated by the other two models.

Unobserved heterogeneity is an important component that should be considered in the modeling process, even though it is not commonly involved in the analysis of the termination risks of the mortgages. In Chapter 3, this study uses latent classes to control unobserved heterogeneity of two different groups of borrowers and constructs three competing risks models based on the multinomial logit, the proportional hazards model and a class of transformation survival models. The models allow the coefficients of the explanatory variables to be different between two groups of borrowers by keeping the baseline the same (the coefficients of the loan age splines are the same between two groups of borrowers). The models are used to analyze the competing risks of prepayment and default/90-days-delinquency of the single-family mortgages in Phoenix and the estimated average conditional hazard for prepayment, default and 90-days-delinquency are compared with those estimated by models that do not control for unobserved heterogeneity. The results show that when the loan age is between 120 and 165 months, models that do not control for unobserved heterogeneity highly overpredict the prepayment hazard. In the average conditional default and 90-days-delinquency hazard, models that do not control for unobserved heterogeneity overpredict the average conditional hazard compared with models that control for unobserved heterogeneity when the loan age is between around 49 and 94 months.

Another question answered in this study is that if housing prices did not boom and bust since 2004, what would the average conditional default hazard and the average conditional 90-days-delinquency hazard be? This paper constructs a simulation process by assuming that the housing price remains the same since September 2004 and compares the simulated conditional hazards with those estimated based on the real trend of the housing price. The results show that, in the case when the housing price changes across time, the average conditional hazard dramatically increases from around age month 11 and reaches the maximum hazard at around age month 54, and then sharply decreases until around age month 93. This dramatic change of the average conditional hazard disappears in the case when the housing price is assumed to be unchanged after September 2004. The simulated average conditional hazard slowly increases from age month 1 up to age month 169 with an average rate of increase of 3.72 percent for default hazard and 1.80 percent for 90-days-delinquency hazard. The average difference of the conditional hazard is approximately 0.21 percent between the two cases.

The Competing Risks of Prepayment and Default on the Single-Family Mortgage Market

by

Ran An

B.S., Capital University of Economics and Business, 2011

M.A., Syracuse University, 2014

Dissertation

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in *Economics*.

Syracuse University

December 2016

Acknowledgements

First, I would like to thank my advisor, Professor Jan Ondrich, for his guidance, support, and hard work. While I was exploring various research topics and unsure of which to focus on, Professor Ondrich provided guidance and focus with his interesting course content and the practical programming advice. He encouraged me to build my own models and trusted that I could solve the complicated calculations by myself. The coursework, readings, and programming advice helped me understand how skilled I am at programming and calculation. After I determined my dissertation topic would focus on the competing risks models in the single-family mortgage market, Professor Ondrich was always available to sit down with me to discuss the ideas to ensure the accuracy of the work. He not only cared about my research achievement but also of my future direction. When I had a great job opportunity in Boston, he fully supported me to start my career before I finished my Ph.D. program. Even after I went to Boston, we remained in touch with each other, sometimes until late into the night, to make sure a dissertation of quality was finished on time. My advisor was like my academic father; someone who fully supported me, trained me with his kindness and patience. I thank him from the bottom of my heart.

Additionally, I extend my gratitude to all of my committee members: Pinyuen Chen, Stuart Rosenthal, John M. Yinger, William C. Horrace, Jeffrey D. Kubik, and Yoonseok Lee. I have sought help from most of them for a long time, both on my studies and future career planning. Without their kindness and help, I would not have successfully defended my dissertation. They raised valuable questions and solutions regarding my dissertation, all of which will be helpful in my career and will be considered in future study.

I also want to thank all of the professors who taught my major courses and for whom I worked with as a teaching assistant. Badi H. Baltagi, Donald H. Dutkowsky, Jerry Evensky, Susan Gensemar, William C. Horrace, Emil Iantchev, Jerry S. Kelly, Jeffrey D. Kubik, Chihwa Kao, Derek Laing, Yoonseok Lee, Devashish Mitra, Piyusha Mutreja, Jan Ondrich, Stuart Rosenthal, Christopher Rohlfs, Perry Singleton, Jeffery Weinstein, and John M. Yinger. They gave me a solid knowledge based in economics, trained me to confidently give presentations in front of audiences and offered me great opportunities to teach and interact with students. All of these experiences are valuable memories of my Ph.D. program. I also want to thank Derek Laing who accepted me into the program five years ago and opened the door toward my achievement.

Further thanks go to Kerr Sougata, Paul Hoffman, and Pruneda Fuentes Alfonso who accepted me to my current position conducting the work I love and enjoy. I also thank my manager, Pal Santanu, who continues to teach me valuable knowledge regarding the banking system, shows me the way to clearly express the ideas and to communicate with people from different backgrounds, and reminds me of the importance of connecting with different business partners. These guidances made my transfer from a student to a professional analyst much smoother. Furthermore, many thanks to my coworkers, Yongmei Shen, Guo Grace, Zizhen Wu, Bose Samrat and Lin Ao. We work together, learn from and help each other to achieve our goals. The spirit of teamwork they provide me has become a valuable new skill.

Moreover, I am truly grateful for my parents, Jianjun An and Hui Huang, who are a teacher and nurse, respectively; they have served in China working for child and teenager development for more than thirty years. As the only child in the family, I grew up with all of their love and support. I can still remember them reading me stories when I was a child every day before I went to bed, competing with me to see who could solve the math problems first,

staying up with me until well into the night when I studied and prepared for my college entrance exam, and woke up early to cook breakfast for me every morning throughout my high school years. They raised me to be a voracious reader, encouraging me to not limit myself to school textbooks but to also read other history, biographies and fiction books. They also provided me with life experiences such as visiting modern cities as well as barren regions to ensure I grew up with independent and reasonable thoughts. Most especially, I want to thank them for supporting my decision to study abroad and further develop the acdemic area I love.

In addition, my eternal gratitude goes out to my grandfather, who treated me fairly with his love. He was the first family member who unconditionally supported my decision to go to the United States to study because of his experiences traveling around the world. Although he passed away three years ago, he will live in my heart forever.

I also extend a great deal of thanks to my Ballroom dance instructor, Dimitry Karabanov, who helped me rebuild my confidence and rearrange my life, enabling me to become an independent, confident woman, thereby enabling me to meet my soul mate, Yang Qiu.

Last but not the least, I send my heartfelt thanks to Yang Qiu, who shares all of my happiness and sorrows, loves me, trusts me, and treats me as his first priority in life.

To all listed above, your continued support and encouragement helped me completed my studies. This Ph.D. degree is the first great achievement in my life, ending twenty-one years of being a hard working student. To quote my manager: “I hope this is the first of many big achievements”, I continue this exciting journey as a risk analyst and with enthusiasm, diligence and great effort.

Contents

Chapter 1

Introduction

The U.S. subprime mortgage crisis, triggered by a large amount of mortgage delinquencies and defaults, draws economists' attention to the termination risks of single family mortgages.

Between 1994 and 2003, when the homeownership rate began a dramatic increase, the subprime share was still relatively stable at around 5.5 percent of mortgage originations (Joint Center for Housing Studies of Harvard University, 2008). However, the subprime share increased rapidly to approximately 20 percent from 2004 to 2006. The financial situation was worsened by the fact a high percentage of these subprime mortgages (approximately 90% in 2006) were adjustable-rate mortgages (ARMs), which were historically offered to the borrowers with good credit, bigger down payments and stable incomes (Zandi Mark, 2009). Although the ARMs were introduced as early as the 1980s, new versions of the ARMs came with extraordinarily low initial rates, known as teasers. At times, teasers were offered at rates of only one or two percent and were fixed for two years in most cases. After two years, they quickly adjusted higher, usually every six months, until they matched higher prevailing interest rates.

In truth, most banks expected these mortgages to be kept for a short period. The expectation was not that these mortgages would default, but that they would be refinanced before the teasers adjusted higher. Most banks allowed borrowers to prepay mortgages with a prepayment penalty, which was usually about 80 percent of the six months' interest. The basic assumption was that, with the strong rise in housing prices, the amount of the second mortgages against the housing equity would increase. This amount of money would be enough to pay back the existing mortgages and the prepayment penalties. Therefore, borrowers could keep the second mortgages with low teasers. For borrowers who have a fixed-rate mortgage, when the market interest rate decreases, they have the incentive to refinance and get a new mortgage with a lower interest rate.

However, housing prices peaked in spring 2006 and then began falling. This meant refinancing was no longer an option because the second mortgages were not enough to pay back the previous ones. Thus, millions of subprime mortgages holders began defaulting on their mortgages and banks foreclosed on their houses. More than four million completed foreclosures entered the market from January 2007 to December 2011, and approximately 8.2 million foreclosures were started (Bennett, 2012). The housing prices dropped dramatically because of a large amount of housing availability and the shortage of housing demands, which triggered even more defaults.

Banks that lent subprime mortgages became undercapitalized because, in many cases, the values of the houses were less than the original mortgages they lent. The low equity-asset ratio and the high loan-asset ratio (which is the amount of properties in the foreclosure process or were foreclosed as a percent of assets) caused 466 banks to be shut down by the Federal Deposit Insurance Corporation (FDIC) from September 2007 to December 2012 (Boswel, 2013). Approximately 7177 banks were still operating as of September 2012. These banks were separated into three groups - the Big Four Banks (Bank of America, JPMorgan Chase, Citigroup and Wells Fargo), the Next Thirteen Largest Banks (US Bancorp, Capital One, PNC, Mellon, State Street, TD Bank, HSBC, etc) and the Mid-and-Small Banks (all the remaining 7160 mid-and-small sized banks). The bank data offered by FDIC showed that the amount of loss provisions as a percent of loans during the pre-crisis period for the Big Four Banks was 2.26%, for the Next Thirteen Largest Banks was 1.31% and for the Mid-and-Small Banks was 0.94%. However, during the crisis period, the amount of loss provisions as a percent of loans for the Big Four Banks, the Next Thirteen Largest Banks the Mid-and-Small Banks increased to 12.38%, 9.30%, and 8.65%, respectively. Since December 31, 2007, approximately $779 billion in loan loss provisions were recorded on US banks income statements. Of that amount, $691 billion was written off in actual Net Charge-offs (FDIC Bank Data) and the amount the Loan Loss Allowance has increased to $167 billion since December 2017.

Another important financial system negatively affected by the delinquencies and defaults during the crisis was the secondary mortgage market. In this market, existing mortgages were sold by lenders and packaged into mortgage-backed securities (MBS), and then bought by public and private investors, including Fannie Mae, Freddie Mac, pension funds, insurance companies, mutual funds and hedge funds ("Mortgage-backed security", 2016). The value of MBS insurance was around 1.35 trillion in 2001 (according to Sifma Statistics), increased to around 2.7 trillion in 2003 and then dropped to around 2.1 trillion in 2006 (gross U.S. insurance agency MBS plus private-label MBS). However, the securitization share of subprime mortgages increased from 54 percent in 2001 to 76 percent in 2004 and then dropped slightly to 75 percent in 2006. Therefore, a large number of subprime mortgages were packaged into MBS'. The inability to predict the value of MBS because of the large rate of defaults trapped many sellers and investors in the secondary mortgage market.

As an important trigger of the subprime crisis, the default risk is worth comprehensive analysis. In addition, the prepayment risk, which greatly influenced the level of the cash flow in the mortgage market, should also be considered. Previous literature has provided evidence that the prepayment and default risks are distinct but not independent. Therefore, when the probability of prepayment and default are simultaneously high, it is interesting to analyze the following research questions: what economic factors affect the competing risks of prepayment and default and to what extent they affect the risks?

To understand these questions, this paper analyzes the economic factors affecting these two competing risks in the single-family mortgage market in locations where the aggregate of the prepayment and default risks are simultaneously high. The model used is a discrete-time competing-risks model based on restricted multinomial logit analyzing information from a sample single-family loan-level dataset constructed by Freddie Mac, which contains thirty-year, fixed-rate mortgages originated from January 1999 to December 2012. Individual mortgages are analyzed according to high simultaneous prepayment and default rates in five MSAs – Phoenix, Miami, Tampa, Detroit and Las Vegas.

To develop a more comprehensive analysis, two different combinations of dependent variables are used. The first combination is prepayment and default and the second is prepayment and 90-days-delinquency. The indicator for a prepayment penalty and the value of the call option[1] are used to evaluate the prepayment risk and the best three combinations of explanatory variables – the value of negative equity, a negative equity dummy and original loan-to-value– together with the unemployment rate are used to evaluate the default risk.

This research also examines the effect of the credit score and the debt-to-income ratio on the termination risks. According to the literature, the credit score affects the probability of default; however, whether or not it affects prepayment is ambiguous. Therefore, two estimations

[1] A call option is an agreement that gives an investor the right to buy a stock, bond, commodity or other instrument at a specified price within a specific time period. Prepayment gives a borrower the right to early repay a loan to take advantage of lower interest rates, because when principal is prepaid early, future interest payments at a higher interest rate will not be paid on that part of the principal. Therefore, in the case of a mortgage-backed security, prepayment is treated as a call option. The value of the call option is the difference between the market value of the mortgage and the book value of the mortgage measured as a percent of the market value of the mortgage.

are run to explain the effect of the credit score on the prepayment and default. The first estimation considers how the credit score affects both termination risks and the second considers how the credit score affects only default. In addition, other mortgage characteristics, such as loan size and loan age, are controlled to make the estimation accurate.

The brief summary of the main results shows the prepayment penalty has a negative effect on the prepayment decision. However, its effect in Detroit is insignificantly positive. The value of the call option has a significantly positive impact on the prepayment in each MSA. The unemployment rate has a positive relationship with the default/90-days-delinquency. Moreover, negative equity, the negative equity dummy, and original loan-to-value generally are positively related to the default/90-days-delinquency. As expected, in most MSAs, the credit score has a strongly positive effect on the prepayment; comparatively, it has a strong negative impact on the default. When using the credit score only in the default risk and not in the prepayment risk, the effect of the credit score is still significantly negative, but the coefficients slightly decrease. Moreover, the debt-to-income does not appear to affect prepayment, but it has a positive effect on the default/90-days-delinquency in most MSAs. To test the equality of the coefficients among five MSAs, a Wald test is constructed in this paper. The results of this test support the argument that there is not a national financial market for housing and that the financial markets for housing are heavily influenced by local characteristics.

The paper is separated into six sections. In section II, previous literature on the analysis of prepayment and default in mortgage markets is summarized. In section III, the dataset, explanatory variables and dependent variables used in this paper are clearly discussed. Section IV outlines the methodology and section V discusses the results of the five MSAs and the Wald test for the equality of coefficients among MSAs. The paper concludes and recommendations for future study are covered in section VI.

Literature Review

In the earliest studies using option models in real estate finance, researchers analyze either prepayment or default decision, but not both. Findley and Capozza (1977) provide evidence that the refinancing risk of variable-rate mortgages (VRMs) and the refinancing risk of fixed-rate mortgages (FRMs) can be reasonably explained by option theory. The analysis only considers interest rate intermediation and by employing the Samuelson-Kruizinga notation, the authors clearly show the effect of the bond price (the reciprocal of the interest-rate level) on key variables, including the market value of a mortgage, an S&L deposit, equity value, gross wealth and net wealth. The analysis is discussed from the separate perspectives of savings and loan associations (S&Ls) and mortgage holders. From the S&L perspective, Findley and Capozza construct a net equity position in terms of the bond price movement for the true FRMs, the actual FRMs and the VRMs. The difference between true FRMs and the actual FRMs assumed by authors is that the former cannot be refinanced or accelerated and the latter can be refinanced by the borrower and accelerated by the lender. Using these net equity positions, the authors argue that the S&L solvency problem consists of forecasting and pricing issues, and the pricing issues include an option pricing problem in the case of the actual FRM. Moreover, both problems can be mitigated by VRMs. From the mortgages holders’ perspective, the authors construct a net wealth equation in terms of the interest rate change for the true FRMs, the actual FRMs and the VRMs. The equations are separately discussed under the assumptions of fixed income, variable income and bounded income. The results show the VRMs have no adverse effect on the households sector. Therefore, allowing S&L to market residential VRMs would be a wise solution for solving the S&L solvency problems.

While the paper by Findley and Cappoza presents a theoretical analysis only, Epperson, Kau, Keenan and Muller's 1985 study is the first that uses simulation to understand default behavior. Epperson et al. analyze the default option and evaluate the pricing of insurance against default on FRMs. In this study, the mortgages and insurance are treated as compound put options and the prepayment option is ignored. A compound put option means that, at each payment date before the last one, the borrowers either default or purchase a new option to default at the next payment date by making the scheduled payment. By assuming the value of the house follows the standard lognormal process and the interest rate follows a stochastic process under the Local Expectations Hypothesis (LEH), a partial differential equation (PDE) is constructed. The valuation of derivative assets as the solution of the PDE involves both housing prices and the interest rate. The simulation process provides three important results. First, the value of the default option increases as the loan-to-value ratio increases. Second, the volatility of both the house price and the spot interest rate significantly affect the value of the default option. However, the effect of the house price is stronger than that of the spot interest rate. Third, when the insurance coverage is 100 percent of the loan and with the increase of the loan-to-value ratio, the percentage increase of the value of the option is less than that of the value of the insurance. This result indicates that the additional insurance coverage is less valuable. Furthermore, the authors recommend that a complete analysis of FRMs should involve both the prepayment and the default option, and the techniques used in this paper can be applied in the future study.

In responding to the gap in empirical analysis, Green and Shoven (1986) collect 3938 fixed-rate mortgages from two large California S&Ls and apply the proportional hazards model to estimate the sensitivity of mortgage prepayments to interest rates. The data is available from 1975 to 1982 for the first association and from 1976 to 1982 for the second. The sample is divided into two periods, 1975-78 and 1979-82 according to whether due-on-sale (DOS) clauses are enforced or not. The DOS gives the lender the right to claim the face value of the mortgage if the borrower sells the residence. The results indicate that the probability of prepayment is sensitive to the relationship between the contract rate and the market rate. Moreover, the effect of the interest rate during the period for which the DOS clause is unenforceable is much higher than that in the period for which it is enforceable. Furthermore, the authors test the sensitivity of the average mortgage age at the time of prepayment to the interest rates and found that the average age was highly dependent on interest rates.

Schwart and Torous (1989) follow Green and Shoven's methodology and add lagged refinancing rates, heterogeneity of borrowers and seasonality into the model to predict the probability of prepayment. Next, they integrate the prepayment function into a valuation framework to evaluate the mortgage-backed security. The dataset used in their study is a number of thirty-year, fixed-rate, single-family pools offered by the Government National Mortgage Association (GNMA) from January 1978 to November 1987. The results show that the conditional probability of prepayment significantly increases when the refinancing rate is less than the contract rate and insignificantly increases during the summer. Furthermore, the conditional probability of prepayment is positively correlated with the mortgage age until around 6.265 years, after which point the relationship between the probability of conditional prepayment and mortgage age becomes negative.

Quigley and Van Order (1990) estimate a prepayment function based on a hazard model and test whether borrowers exercise their prepayment options in a manner consistent with the optimal strategy developed in contingent claims models. The dataset is from the Federal Home Loan Mortgage Corporation (Freddie Mac) and includes 6375 thirty-year fixed-rate single-family mortgages issued during the 1976-1980 period. The results show the value of the call (prepayment) option being "in the money" has a strongly positive effect on the prepayment hazard, and the initial equity insignificantly and negatively affects the prepayment hazard. However, after testing the prepayment behavior, the authors conclude the prepayment option is not exercised ruthlessly. In their paper, the authors also mention they do not consider default in the choice model; however, a more general model should consider that loans are about to prepay as well as to default.

Follain, Ondrich and Sinha (1995) explain that the substantial transaction costs, nonfinancial characteristics – such as job change and divorce – personal desire to adjust the composition of portfolio, and data limitation are four main reasons that homeowners do not ruthlessly make the prepayment decision. Therefore, the authors argue the theoretical predictions of the option pricing model do not hold empirically on single family mortgages. In addition, they introduce a prepayment model based on the proportional hazard model developed by Meyer (1987) to explain the prepayment behavior in the multifamily mortgage market. They also consider unobserved heterogeneity in this model. The dataset contains 1083 Freddie Mac Plan A multifamily mortgages originated between 1975 and 1986. An important result in their study is that the prepayment is positively correlated with the value of the call option. When the call option is in the money, households are more likely to prepay their mortgages. Significant increases in the model with heterogeneity. However, the sensitivity of the prepayment with respect to an interest rate change for an in-the-money option is much lower that what is expected in the option price model. Therefore, the prepayment option is not exercised ruthlessly in the multifamily mortgage market. Moreover, the survival time positively affects the prepayment hazard, and the seasonal effect for summer significantly lowers the prepayment hazard compared to winter. The authors also discuss the effect of the logarithms of housing and book values in the study. In the model without heterogeneity, the effect of the logarithms of housing value is insignificantly positive and the effect of the logarithms of book value is insignificantly negative. However, both effects become significant in the model with heterogeneity. Furthermore, the authors also mention the potential importance of incorporating the value of the default option in the prepayment model.

Quigley and Van Order (1991) estimate the probability of default in different LTV level and in different geographic regions based on 300,000 thirty-year, fixed-rate conventional loans originated from 1976 through 1980 and bought by the Federal Home Loan Mortgage Corporation (Freddie Mac). The data is separated into 15 groups based on three loan-to-value (LTV) categories (LTV is less than 81 percent, between 81 and 90 percent, and greater than 90 percent) and five geographic regions (Northeast, North-central, Southeast, Southwest and West). By using a proportional hazard model with the baseline hazard varies by age, the authors show that a mortgage institution lending loans with LTV above 90 percentage needs about two times of capital more than the one lending loans with LTV between 81 and 90 percentage. Holding the LTV constant, a nationally-diversified institution needs about one-half the capital less than the one located entirely in one region. Using the same dataset, Quigley and Van Order (1995) test the contingent claims approaches for “ruthless” default (the transactions costs, reputation costs and moving costs have no effect in the model) and estimate the relationship between the value of equity and the default risk. The results show that the probability of default increases as the value of negative equity increases. The loans with a higher value of negative equity are more likely to default and loans with a lower value of negative equity do not default immediately. Moreover, by adding the transaction cost into the model, the authors present that it significantly affects the default decision and should be considered in the estimation.

In responding to the fact the prepayment risk and the default risk should be analyzed simultaneously, Kau, Keenan, Muller and Epperson (1992, 1995) write a series of papers using a generalized valuation model, which involves both risks for thirty-year fixed-rate single-family mortgages. The model, in the paper published in 1992, is based on the partial differential equation (PDE) introduced by Epperson *et al*. in 1985. However, the new PDE consider the probability of prepayment occurring at any time and the probability of default occurring only at the end of each month. The factors used in the 1992 study are separated into two groups. The first group is financial factors: interest rate, housing price volatility, and loan-to-value ratio. Nonfinancial factors are included in the second group: divorce and job change. This simulation shows that the lower housing price volatility and the lower loan-to-value ratio has a significant effect on the default decision, while the high loan-to-value ratio and high house price volatility have a significant effect on prepayment. When insurance is added to mortgages, the nonfinancial factors affected the prepayment and default decisions differently. This means borrowers prepay mortgages for nonfinancial reasons; however, they hardly default for nonfinancial reasons. In the paper published in 1995, the authors continue to develop the general valuation model under the assumption of perfect financial markets. The simulation is more general and considers as many of the features of the mortgage contract as possible. However, the general valuation models in both studies are not applied to the actual dataset.

In more recent empirical studies, the Cox proportional hazard framework has become a popular method to construct competing risks models. The basic assumption is that borrowers make the prepayment or default decision on the basis of market conditions to maximize net wealth. Deng, Quigley and Van Order (1996) use the competing risks model to analyze the thirty-year fixed-rate single-family mortgages. The dataset is from Freddie Mac and contains 780,443 mortgages issued from 1976 to 1983 in 26 major metropolitan areas. The important explanatory variables used in this study include the value of the call option (calculated as the ratio of the present discounted value of the unpaid mortgage balance at the current quarterly market interest rate relative to the value discounted at the contract interest rate), the probability of negative equity, the income ratio (calculated as the ratio of housed hold reported income to the MSA median income level), the unemployment rate, and the annual divorce rate. The results show that the value of the call option has a significantly positive effect on the prepayment hazard. Similarly, a higher probability of negative equity positively affects the default hazard. However, a higher probability of negative equity decreases the prepayment hazard. Moreover, lower income borrowers are more likely to default than higher income borrowers. Furthermore, the higher the unemployment rate and divorce rate, the lower the prepayment hazard. In this study, the authors also use a simulation method to estimate the subsidy provided, and the program costs, of zero-downpayment mortgages. The estimation shows that the zero-downpayment loans cost much more than the loans with 5 and 10 percent down-payments.

The competing risks model based on the Cox proportional hazard framework is developed further in later studies. A paper by Deng (1997), introduces a competing risks model in which a binomial mean-reverting interest rate model is added. This model can be used in situations where transactions costs in the mortgage termination are not zero and perfect information on future interest rate movement is not obtained. The dataset used in this paper is from Freddie Mac and contains 489,372 thirty-year fixed-rate single-family mortgages issued in

30 major metropolitan areas from 1976 to 1983. The study concludes that the probability of negative equity has a significantly positive effect on the default risk; however, its effect on the prepayment risk is significantly negative. Similarly, the call option is strongly correlated with the prepayment risk. Moreover, the initial loan-to-value ratio is positively and significantly associated with the default risk. Furthermore, higher unemployment rates increase the probability of default and decrease the probability of prepayment. When comparing the results from the competing risks model with and without a binomial mean-reverting interest rate model, the prepayment behavior becomes less sensitive to the value of the call option in the model with a binomial mean-reverting interest rate model. Moreover, in this model, the negative correlation between the prepayment and default functions becomes more significant.

Deng, Quigley, and Van Order (2000) introduce the unobserved heterogeneity into the competing risks model. The dataset is from Freddie Mac and contains 447,042 thirty-year fixed-rate single-family mortgages issued in 30 major metropolitan areas. Their results are consistent with previous studies, where the value of the call option has a positive relationship with the prepayment risk and a higher probability of negative equity increases the default risk and decreases the prepayment risk. However, the marginal effect of the call option increases about 20 percent in the model where unobserved heterogeneity is considered. Moreover, the higher original loan-to-value ratio causes a higher default risk and it only slightly increases the prepayment risk. The unemployment and divorce rates have a significantly positive effect on the default risk. When adding these rates into the model, the estimated heterogeneity variance declines. The results indicate that ignoring heterogeneity among mortgage borrowers may lead to downward biases in estimating the prepayment risk.

Wenyi Huang and Jan Ondrich (2002) use a bivariate Heckman-Singer nonparametric random effects distribution to control the unobserved heterogeneity. The primary goal of this study is to analyze the competing risks of prepayment and default in the multifamily mortgage market by a competing risks model based on the Cox proportional hazard model. The dataset contains 4006 FHA-insured multifamily mortgages originated from 1980 to 1995. The results indicate the value of the call option is positively associated with the prepayment risk. The unexpected growth in households and the rental price growth rate are negatively correlated with the claim risk. However, the vacancy rate has a significantly positive effect on the claim risk. Furthermore, mortgage size has a positive effect on both the prepayment risk and the default risk. When controlling for unobserved heterogeneity, the effects of the unexpected growth in households, rental price growth rate, and vacancy rate on claim risk became much stronger. Moreover, the effect of mortgage size also was greater in the model controlling for heterogeneity. Therefore, heterogeneity among mortgage borrowers should be considered in estimating the competing risks.

A working paper by Ken Lam, Robert M. Dunsky, and Austin Kelly used two different combinations of dependent variables to indicate the competing risks. The first combination is prepayment and foreclosure completion, and the other combination is prepayment and 90-day delinquency. The authors discuss the impact of down-payment underwriting standards on loan performance. The competing risks model is based on a multinomial logit. By controlling for many important personal specific characteristics, such as the credit score and debt-to-income, the authors argue the lifetime delinquency and foreclosure rates increase monotonically and nonlinearly with the original loan-to-value.

Another important topic on mortgage termination risks is the strategic default risk. The strategic default means that even though the borrowers have the ability to pay back the mortgages, they are willing to default since the value of the property falls below the unpaid principal balance. An example of the studies on the strategic mortgage default is by Bradley, Cutts and Liu (2015). Different from the previous literature which ambiguously distinguish the strategic default from other types of default and ignore the importance of negative equity as a necessary condition for strategic default, this paper use a unique dataset from CoreLogic loan-level data and CoreLogic local market data, which contains both mortgage specific and individual specific characteristics, to accurately measure the strategic default based on the value of negative equity and the capability of payment. By using a logistic model, the authors analyze the effect of the local foreclosure rate and the local strategic default rate on borrower's strategic default decision, by controlling individual characteristics, mortgage characteristics, and economic factors. The results show that the local foreclosure rate and the local strategic default rate are statistically significant and economically important determinants of strategic default indicating contagion.

The paper by Chao Ma (2015) claims that there are two types of default generated by two mechanisms: illiquidity-triggered default and strategic default. The illiquidity-triggered default occurs when households do not have enough money to make their monthly mortgage payment and consequently are forced to default. Unemployment can be one of the reasons causes this type of default. The strategic default (or ruthless default) is caused by the financial incentive to default. In this case, compared with prepayment or continuing to pay, default provides households higher utility. The author also argues that there are two types of prepayment, one is caused by the incentive to refinance and the other one is caused by moving. Unfortunately, in mortgage loan performance datasets, prepayment (default) can be easily observed, but it is hard to distinguish two types of prepayment (default) by different mechanisms. To be able to separately estimate the probability of illiquidity-triggered default, the probability of strategic default, the probability of refinancing, and the probability of prepayment cause by moving, the author assumes that, at each period t, the household firstly determines whether to choose illiquidity-triggered default and if illiquidity-triggered default does not occur, then the household decide whether to move and prepay the mortgages. If neither illiquidity nor move occurs in the previous two processes, the household will have the opportunity to make a choice among continuing to pay on the mortgage, refinance or strategic default. The author introduces a two stages estimation based on multinomial logit. In the first stage, the author analyzes the competing risks of prepayment and default by assuming the probability of default is the sum of the probability of illiquidity-triggered default and the probability of strategic default, and the probability of prepayment is the sum of the probability of refinancing and the probability of moving. In the second stage, the author presents a simulation based on the results in the first stage and uses the results of the simulation together with a multinomial logit model to estimate the probability of different types of prepayment and default. Based on 20,487 30-year fixed mortgages offered by the Ohio Housing Finance Agency, the author shows that the county level unemployment rate positively and significantly affects the probability of illiquidity-triggered default. The value of the default option (the difference of current home value and the unpaid balance) has a significant positive effect on the strategic default, and the value of the refinance option (the difference of the market value and the book value of the unpaid balance) has a significant positive effect on refinance.

This study does not separate the illiquidity-triggered default and the strategic default but involves both unemployment rate and the value of negative equity to analyze the default risk. Moreover, this study does not distinguish the prepayment caused by refinancing from the prepayment caused by moving but uses prepayment penalty and the value of the call option to explain the financial incentive to prepay. Future study can be done by separating different types of default and prepayment with a more complicated competing risks model.

Follain and Giertz (2016) use an annual pool time series dataset which includes 384 MSAs from 1980 to 2010 and a three-equation vector correction model to predict the probability of the housing price bubbles. The authors separate the dataset into 9 categories based on the population size (less than 500,000 residence, 500,000 to 1 million residence, and more than 1 million residence) and three time periods (1980-1995, 1980-2000, and 1980-2007) and run regressions on each category. Based on the regression results, the authors predict the housing price from 2008 to 2010 and compare the predicted housing price with the actual housing price. The comparison shows two important results that the model systematically underestimated the sensitivity of the declines and both the predictions and the actual outcomes show dramatic variations among MSAs. Therefore, the authors claim that there is not a national housing market and the housing markets are heavily influenced by local characteristics. Methods that not allow for the variation across markets will cause estimation biases.

Data Description and Explanatory Variables

Data Description

The dataset used in this paper is a sample single-family loan-level dataset constructed by Freddie Mac. This sample dataset[2] is randomly selected from 16 million thirty-year fixed-rate[3] single-family mortgages originated from January 1, 1999 through December 31, 2012[4]. 50,000 mortgages are selected from each year, which amounts to 700,000 mortgages.

To be able to use the Case-Shiller indices for the housing value, the initial analysis in this chapter is restricted to the twenty MSAs for which they are available. Five of these MSAs have high simultaneous prepayment and default rates in the latter part of the sample period. Because the method in this paper involves the competing prepayment and default risks, the concentration is on these five MSAs: Phoenix, Miami, Tampa Bay, Metro Detroit and Las Vegas. Figures 1 through 5 show the empirical prepayment rates and default rates in these MSAs. The calculations of the prepayment and default rates are explained in the Appendix A.

[insert Figures 1 through 5 here]

As the above figures show, the prepayment rates continue to increase from August 1999 to July 2003, and then dramatically decrease until the end of 2008, and then slowly increase again. The default rates are low in each month from 1999 to 2007; however, after March 2008, the default rate substantially increases until March 2011. Thus, in the period from around 2008 to 2011, the prepayment and default rates are simultaneously high for these five MSAs.

---

[2] According to the General User Guide, the sampling method used in this dataset does not change across years, and each member of the subset has an equal probability of being chosen once from the larger population. A simple random sample is meant to be an unbiased representation of the larger population.

[3] Original sample dataset includes any loan with loan term greater than 300 months and less than 420 months. However, the dataset used in this paper contains only thirty-year (loan term is 360 months) fixed-rate single-family mortgages.

[4] The actual originations in the sample dataset range from January, 1999 through March, 2013. Some loans originated in 2013 are included in the 2012 sample dataset.

Figure 1. Monthly prepayment rate and default rate in Phoenix

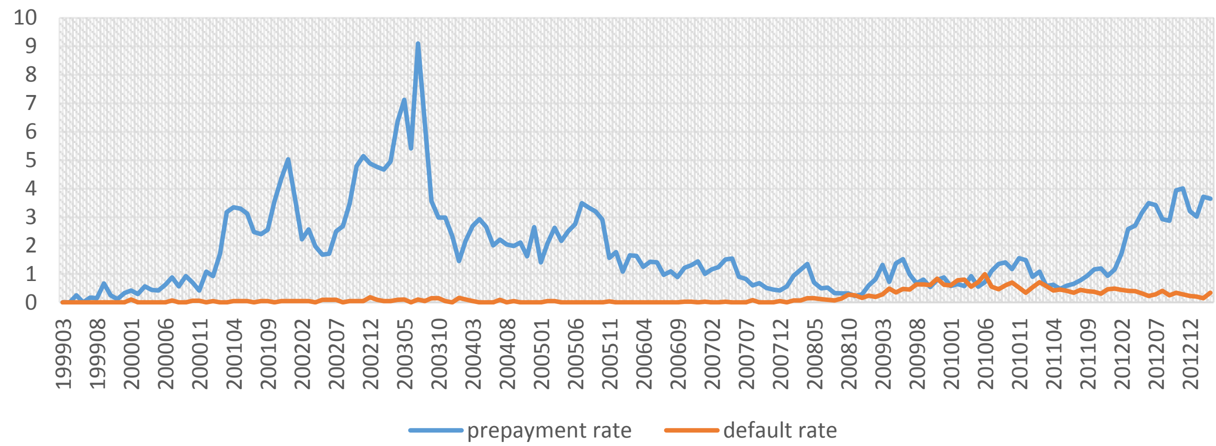


Figure 2. Monthly prepayment rate and default rate in Miami

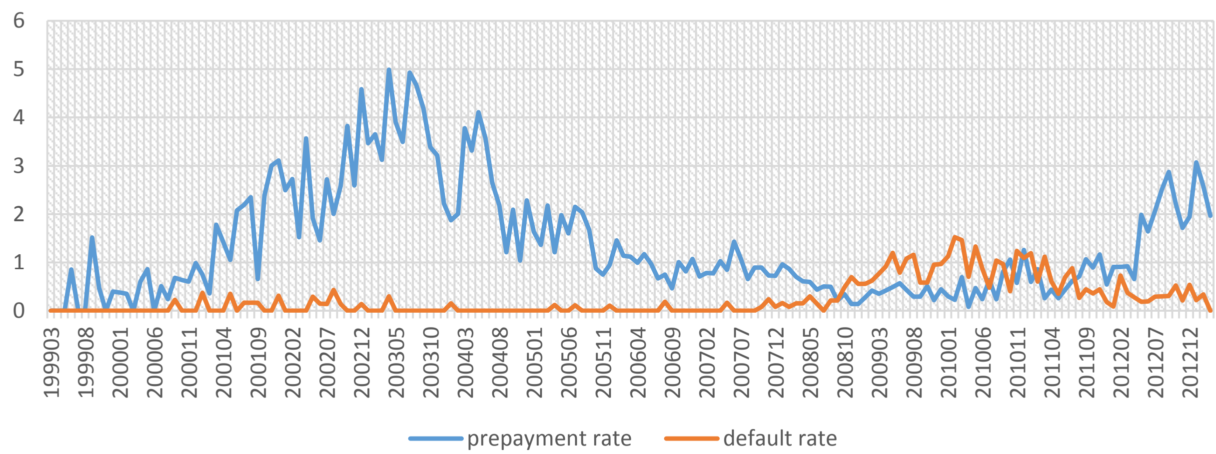


Figure 3. Monthly prepayment rate and default rate in Tampa

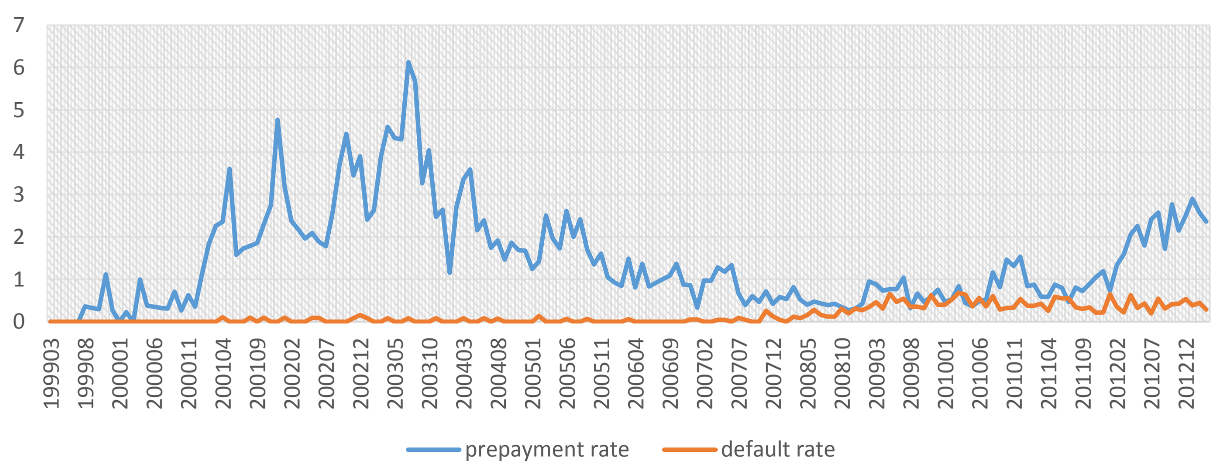


Figure 4. Monthly prepayment rate and default rate in Detroit

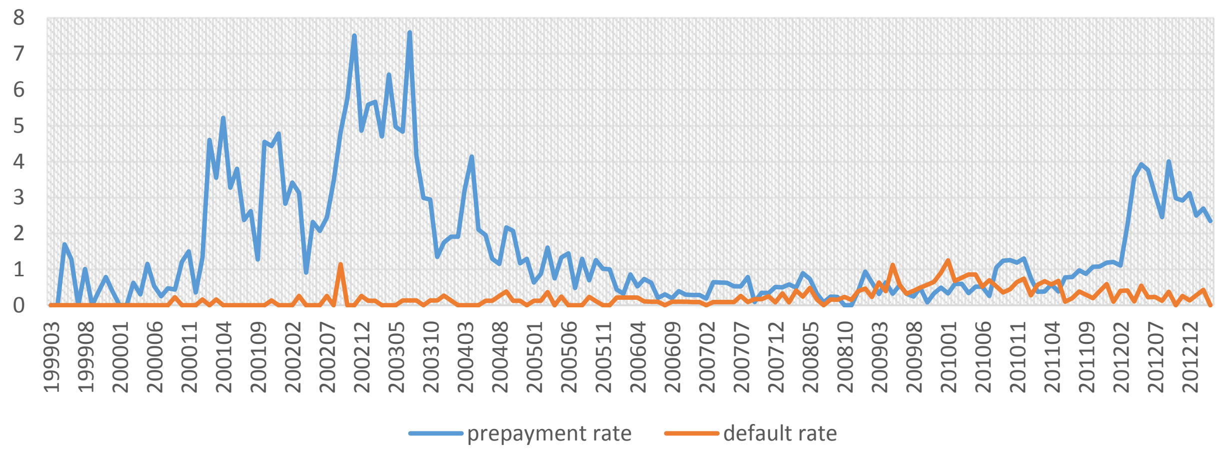


Figure 5. Monthly prepayment rate and default rate in Las Vegas

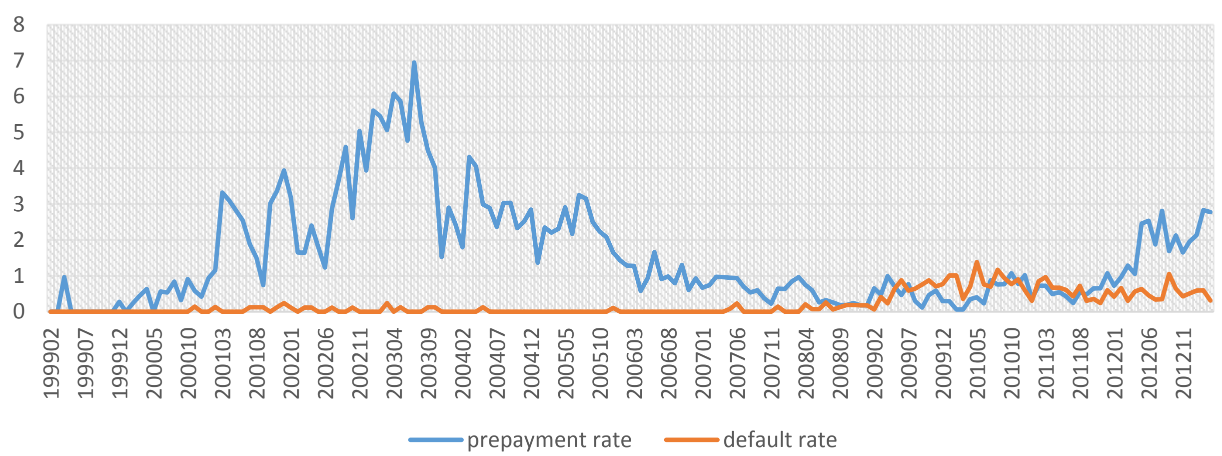


Table 1 shows that a large percentage of mortgages in each MSA has a positive call option and negative equity simultaneously from Aug. 2009 to Jun. 2010. An average of 47.46 percent of existing mortgages in Phoenix, 38.43 percent of existing mortgages in Miami, 39.93 percent of existing mortgages in Tampa, 65.56 percent of existing mortgages in Detroit and 62.02 percent of existing mortgages in Las Vegas with positive call option and negative equity simultaneously. Therefore, the prepayment and default risks in this period are to a large extent simultaneously high. The distribution of the negative equity and call-option for each MSA is shown in Figure 6 through Figure 23 [5]. Table 2 shows the estimated prepayment and default risks for each MSA from Aug. 2009 to Jun. 2010. The results support the argument that there is a competing risk for prepayment and default in the sample period. An average of 10.18 percent of existing mortgages in each MSA has a difference of predicted prepayment and default risks that is less than 0.1 percent. An average of 26.75 percent of existing mortgages in each MSA has a difference of predicted prepayment and default risks that is less than 0.25 percent. An average of 51.60 percent of existing mortgages in each MSA has a difference of predicted prepayment and default risks that is less than 0.5 percent.

[insert Tables 1 and 2 here]

[insert Figures 6 through 23 here]

The findings indicate that 442 mortgages have been modified before the original maturity date. The modification changes the mortgage term, maturity date and contract rate. Therefore, these 442 mortgages will have different monthly payments and remaining balances.

[5] For Phoenix and Miami, the distribution of call option and negative equity is shown for Aug. 2009, Jan. 2010 and Jun. 2010. For Tampa, Detroit and Las Vegas, the distribution of call option and negative equity is only shown for Aug. 2009, because the distributions for other time period are very similar.

Table 1.The percent of mortgages with positive call-option and negative equity in each MSA

Time	Phoenix	Miami	Tampa	Detroit	Las Vegas
Aug. 2009	53.21%	40.52%	38.89%	67.52%	66.14%
Sep. 2009	51.41%	38.80%	39.13%	64.50%	64.42%
Oct. 2009	48.48%	39.10%	39.57%	60.50%	61.40%
Nov. 2009	47.29%	38.54%	40.08%	60.95%	61.22%
Dec. 2009	47.30%	38.94%	40.48%	63.26%	62.33%
Jan. 2010	47.50%	38.71%	40.91%	65.43%	63.16%
Feb. 2010	47.58%	38.26%	42.12%	67.01%	63.02%
Mar. 2010	47.13%	38.77%	41.45%	69.97%	61.76%
Apr. 2010	45.11%	38.71%	40.00%	69.01%	60.50%
May. 2010	43.93%	36.80%	38.70%	67.95%	59.46%
Jun. 2010	43.13%	35.56%	37.91%	65.10%	58.78%

Table 2.The percent of mortgages with the absolute difference of the predicted prepayment and default rates less than 0.1%, 0.1%-0.25% and 0.25%-0.5%

	Phoenix			Miami			Tampa			Detroit			Las Vegas		
	<0.1%	0.1%-0.25%	0.25%-0.5%	<0.1%	0.1%-0.25%	0.25%-0.5%	<0.1%	0.1%-0.25%	0.25%-0.5%	<0.1%	0.1%-0.25%	0.25%-0.5%	<0.1%	0.1%-0.25%	0.25%-0.5%
Aug. 2009	4.4%	12.7%	18.7%	13.3%	17.8%	28.6%	7.7%	14.4%	23.9%	13.6%	17.5%	27.8%	12.8%	19.5%	29.1%
Sep. 2009	6.0%	15.8%	22.0%	15.5%	20.6%	27.3%	9.9%	16.0%	30.1%	15.6%	21.5%	26.0%	14.5%	21.6%	30.6%
Oct. 2009	6.5%	15.0%	23.6%	15.9%	21.2%	25.2%	9.9%	17.4%	29.9%	14.9%	22.1%	24.9%	14.0%	22.5%	29.3%
Nov. 2009	6.2%	14.8%	23.5%	15.3%	21.4%	26.3%	9.9%	17.8%	29.1%	15.0%	21.3%	23.8%	14.8%	21.4%	27.8%
Dec. 2009	4.6%	12.0%	21.7%	13.9%	19.2%	26.4%	8.8%	15.3%	25.8%	11.1%	20.0%	24.1%	11.3%	21.7%	27.9%
Jan. 2010	4.2%	9.7%	19.9%	11.7%	18.0%	27.4%	7.4%	13.9%	21.5%	10.0%	16.7%	24.1%	9.4%	21.0%	26.3%
Feb. 2010	3.9%	9.4%	18.7%	11.0%	16.9%	28.0%	7.0%	12.8%	21.0%	10.2%	14.3%	24.7%	8.9%	19.9%	25.2%
Mar. 2010	4.4%	10.7%	19.0%	11.9%	18.3%	26.3%	8.0%	13.2%	21.9%	12.5%	13.3%	24.8%	11.9%	18.2%	26.5%
Apr. 2010	3.8%	11.0%	18.6%	12.4%	17.6%	27.4%	7.9%	13.1%	22.2%	12.8%	13.9%	24.3%	12.0%	17.4%	27.5%
May. 2010	3.5%	10.6%	18.4%	12.2%	16.1%	29.7%	7.7%	13.0%	21.4%	11.7%	14.3%	23.7%	10.9%	18.4%	25.7%
Jun. 2010	4.1%	11.6%	18.9%	12.3%	18.0%	28.7%	7.9%	13.8%	22.4%	12.4%	15.3%	23.9%	10.3%	20.5%	25.2%

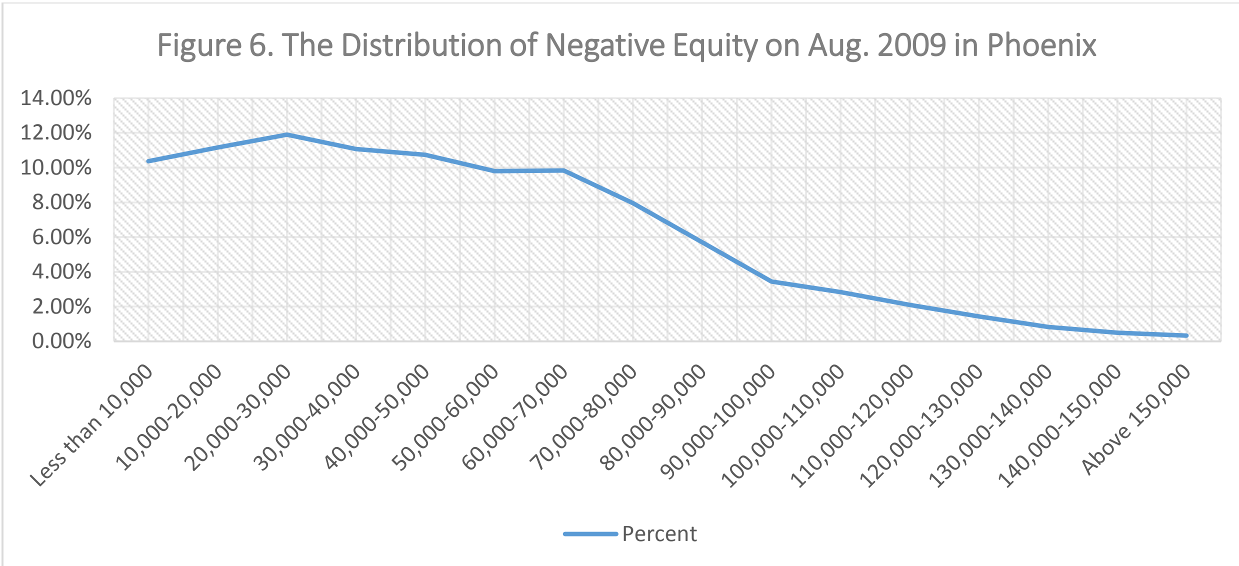


Figure 6. The Distribution of Negative Equity on Aug. 2009 in Phoenix
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
Less than 10,000
10,000-20,000
20,000-30,000
30,000-40,000
40,000-50,000
50,000-60,000
60,000-70,000
70,000-80,000
80,000-90,000
90,000-100,000
100,000-110,000
110,000-120,000
120,000-130,000
130,000-140,000
140,000-150,000
Above 150,000
Percent

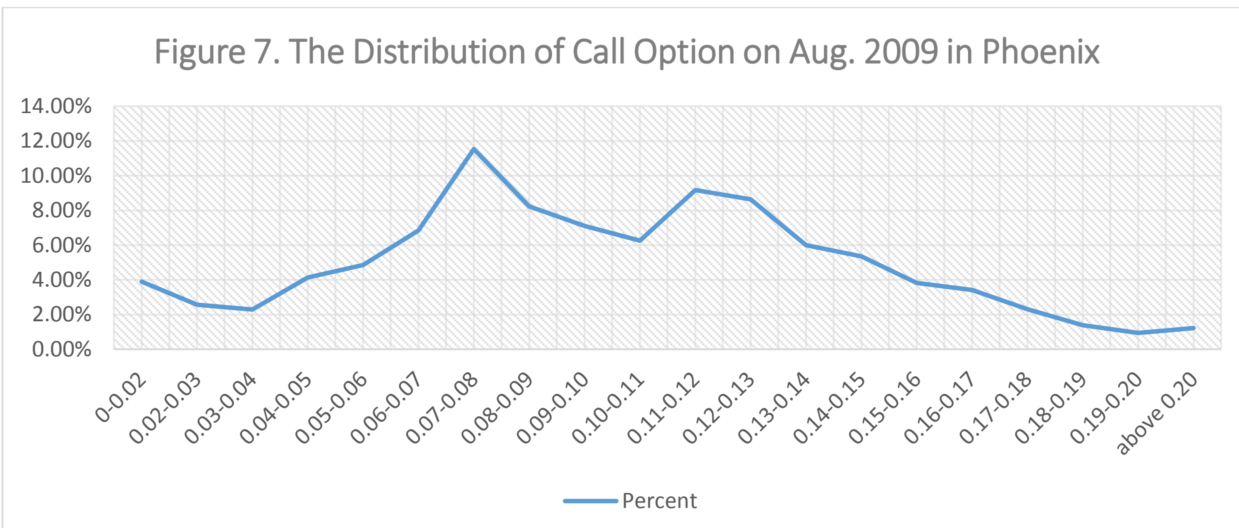


Figure 7. The Distribution of Call Option on Aug. 2009 in Phoenix
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
0-0.02
0.02-0.03
0.03-0.04
0.04-0.05
0.05-0.06
0.06-0.07
0.07-0.08
0.08-0.09
0.09-0.10
0.10-0.11
0.11-0.12
0.12-0.13
0.13-0.14
0.14-0.15
0.15-0.16
0.16-0.17
0.17-0.18
0.18-0.19
0.19-0.20
above 0.20
Percent

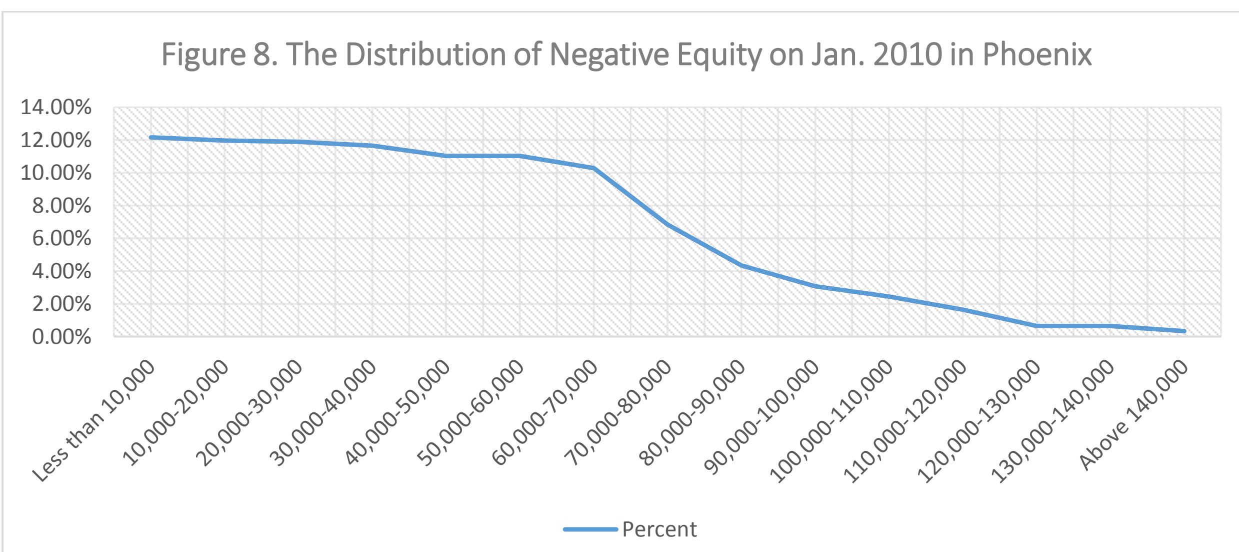


Figure 8. The Distribution of Negative Equity on Jan. 2010 in Phoenix
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
Less than 10,000
10,000-20,000
20,000-30,000
30,000-40,000
40,000-50,000
50,000-60,000
60,000-70,000
70,000-80,000
80,000-90,000
90,000-100,000
100,000-110,000
110,000-120,000
120,000-130,000
130,000-140,000
Above 140,000
Percent

Figure 9. The Distribution of Call Option on Jan. 2010 in Phoenix

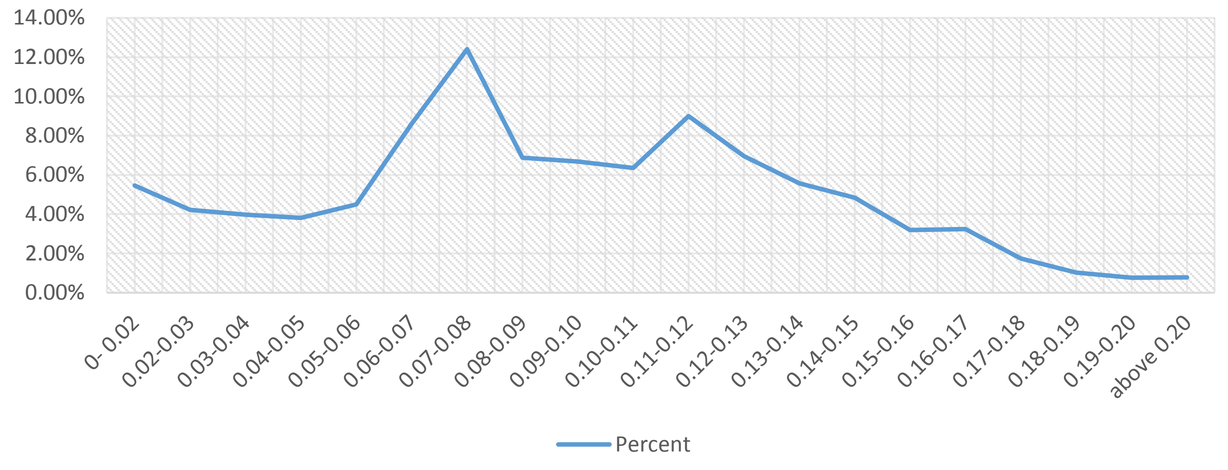


Figure 10. The Distribution of Negative Equity on Jun. 2010 in Phoenix

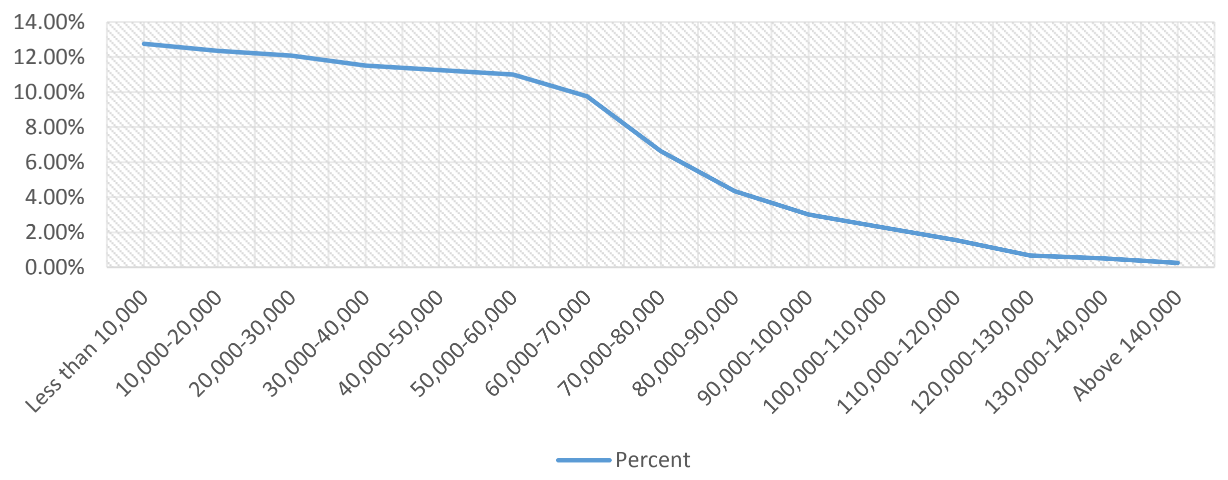


Figure 11. The Distribution of Call Option on Jun. 2010 in Phoenix

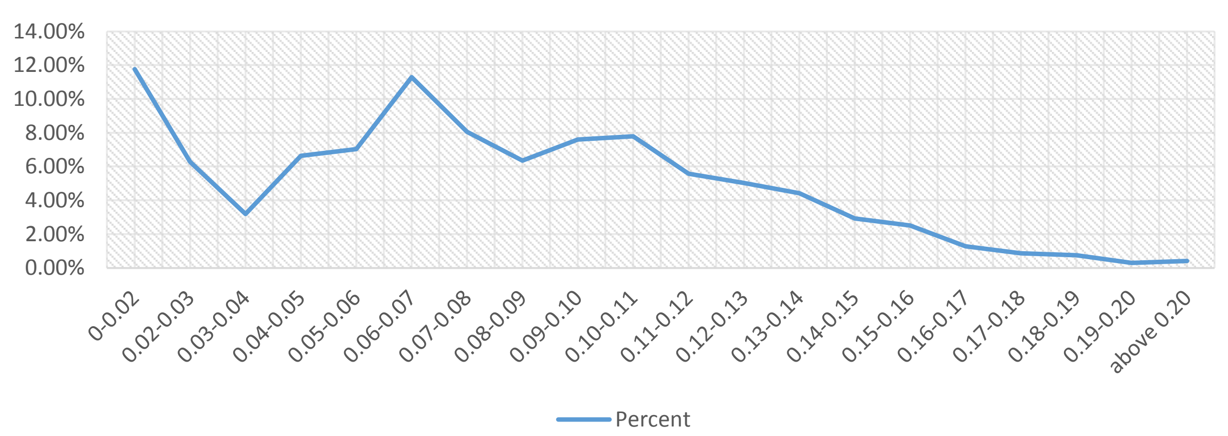


Figure 12. The Distribution of Negative Equity on Aug. 2009 in Miami

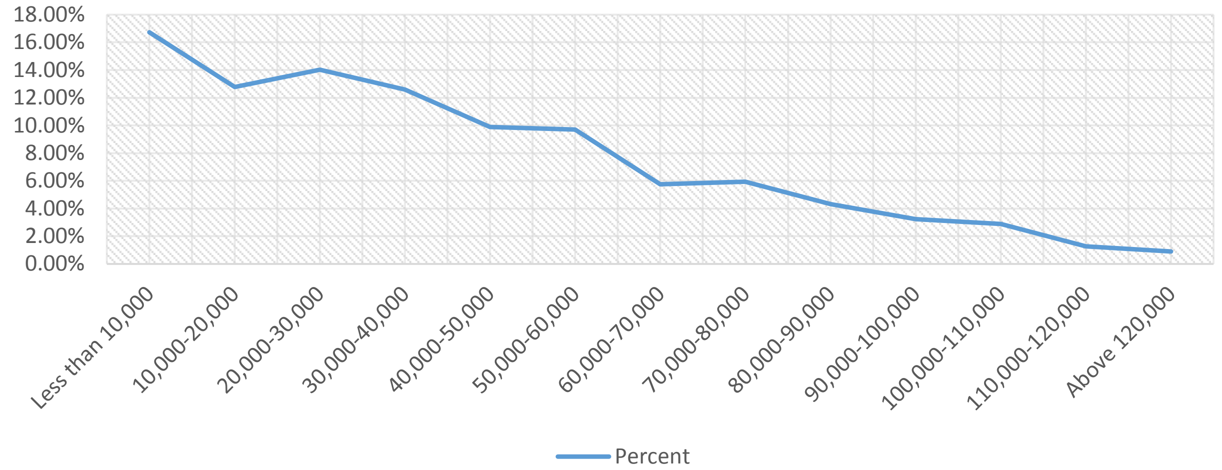


Figure 13. The Distribution of Call Option on Aug. 2009 in Miami

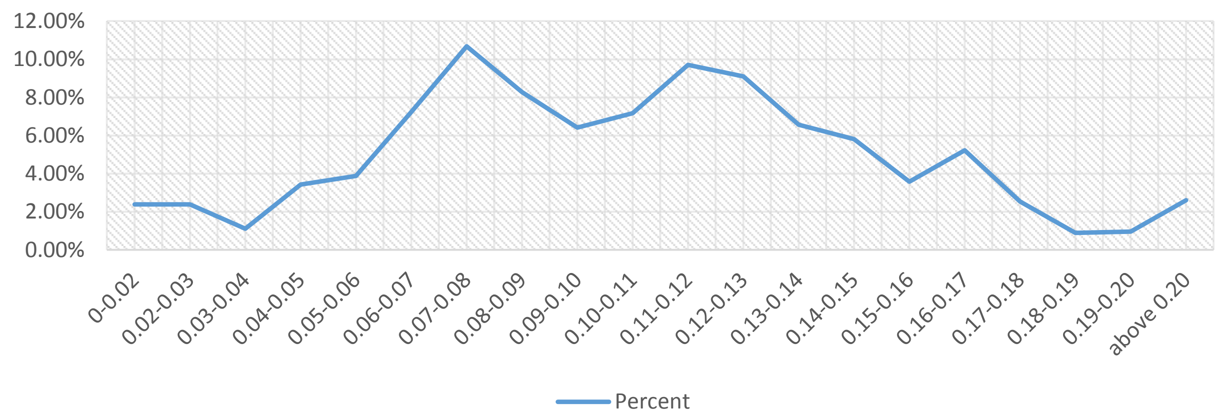


Figure 14. The Distribution of Negative Equity on Jan. 2010 in Miami

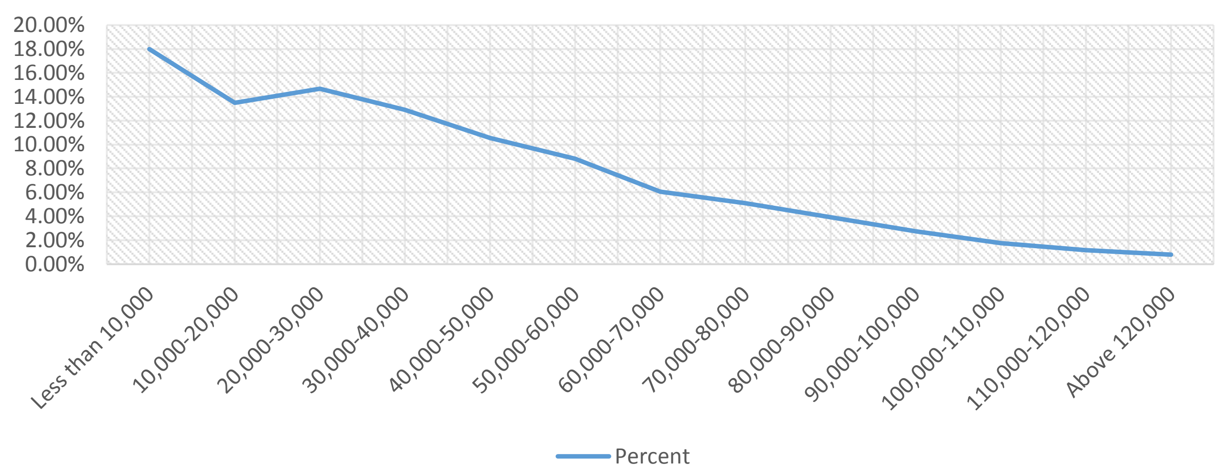


Figure 15. The Distribution of Call Option on Jan. 2010 in Miami

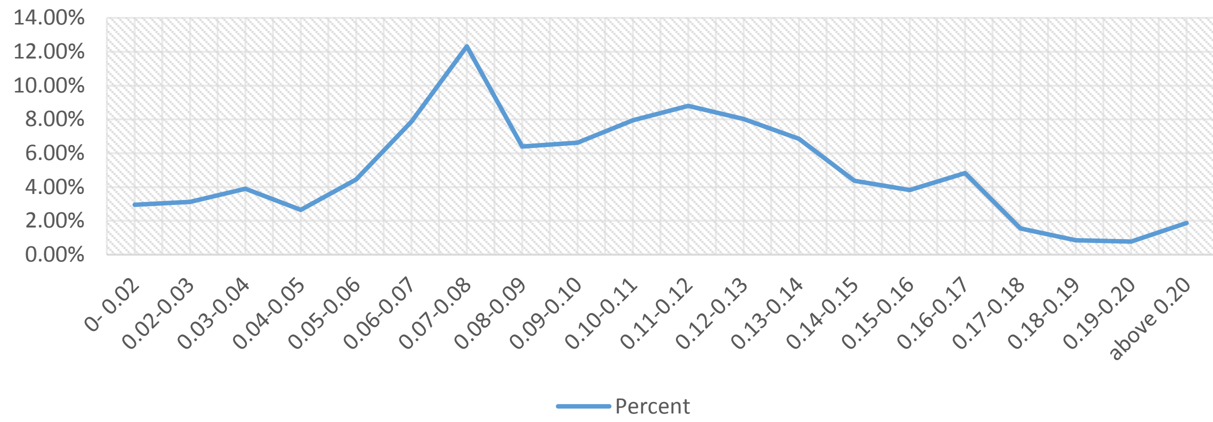


Figure 16. The Distribution of Negative Equity on Jun. 2010 in Miami

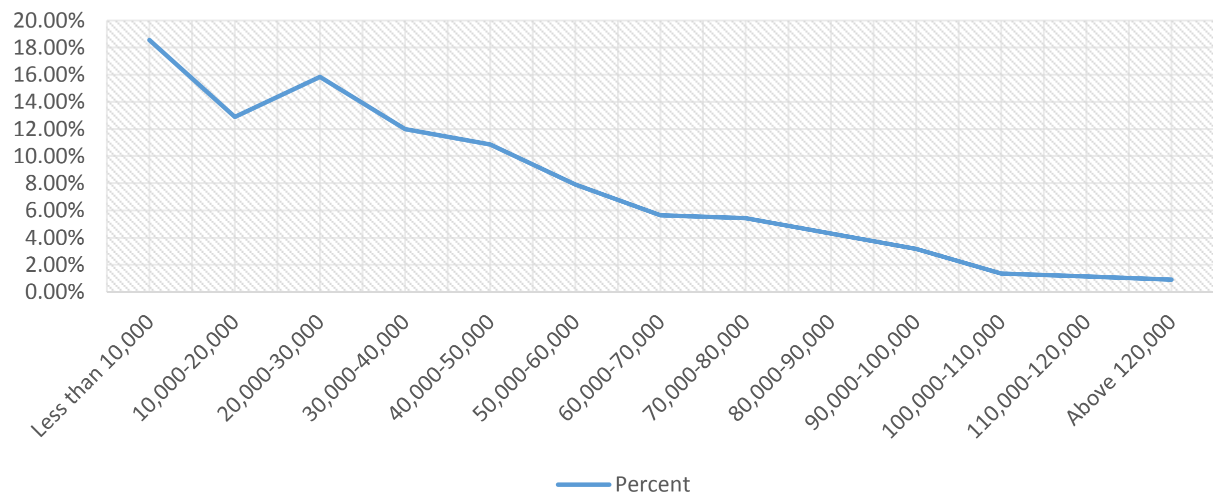


Figure 17. The Distribution of Call Option on Jun. 2010 in Miami

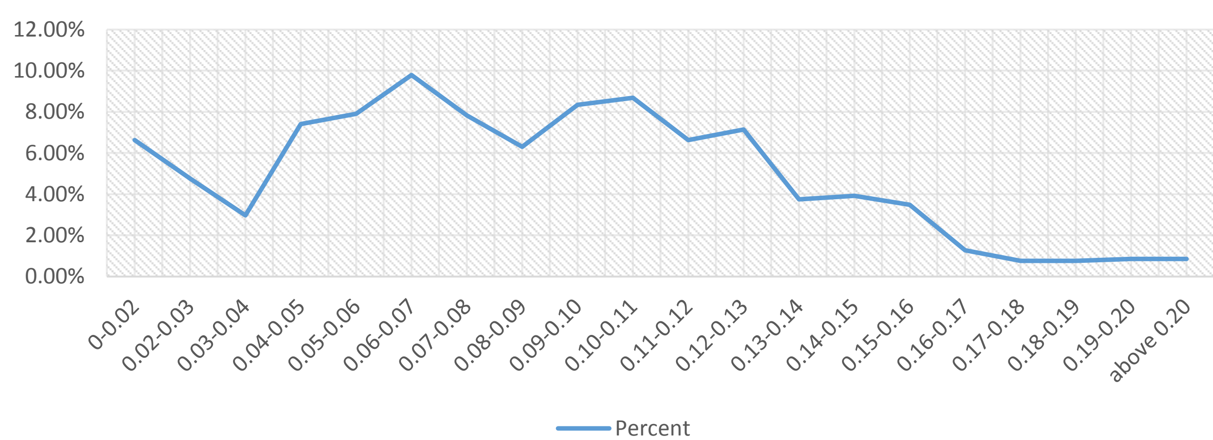


Figure 18. The Distribution of Negative Equity on Aug. 2009 in Tampa

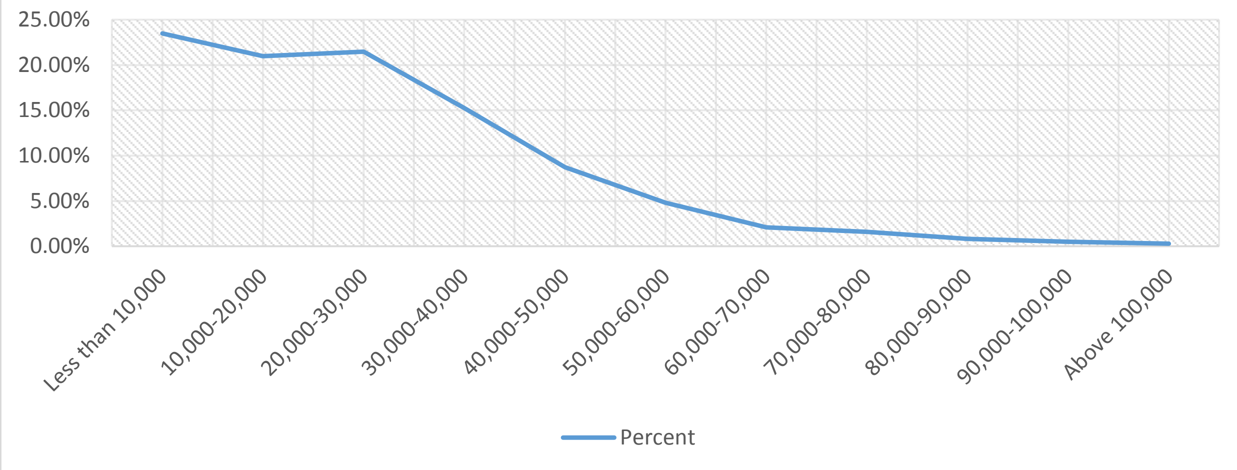


Figure 19. The Distribution of Call Option on Aug. 2009 in Tampa

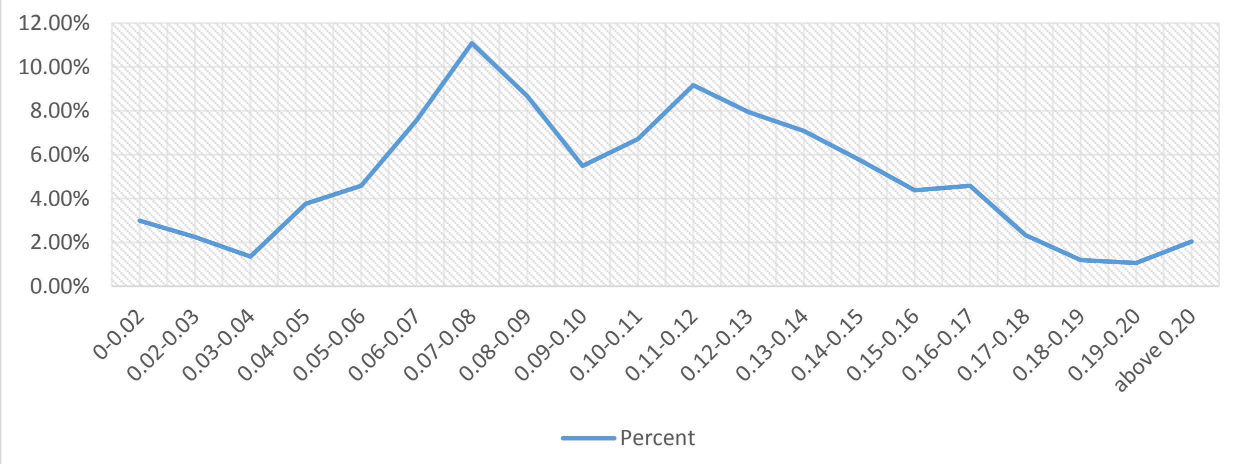
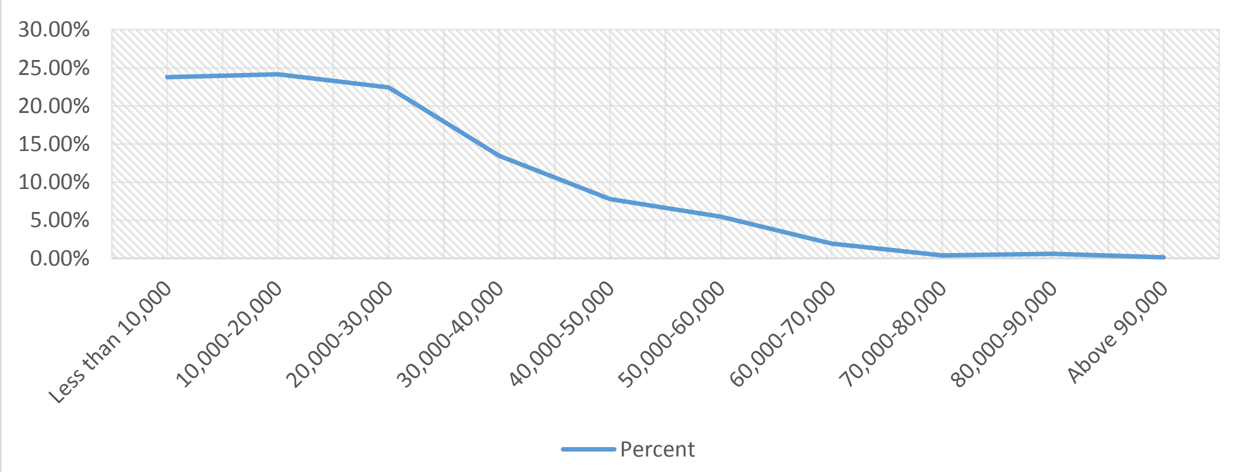


Figure 20. The Distribution of Negative Equity on Aug. 2009 in Detroit



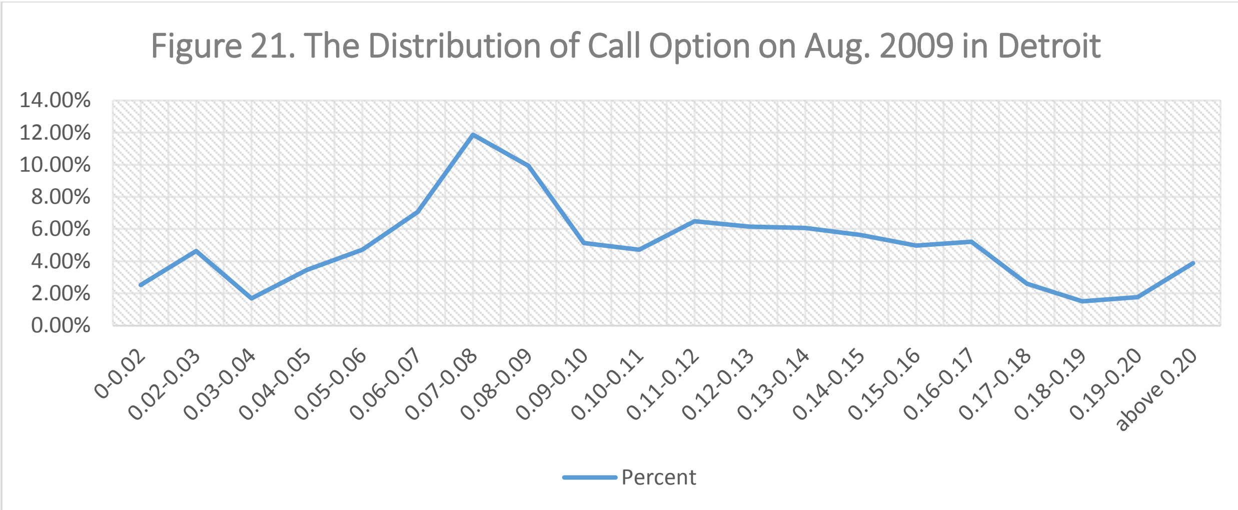


Figure 21. The Distribution of Call Option on Aug. 2009 in Detroit
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
0-0.02
0.02-0.03
0.03-0.04
0.04-0.05
0.05-0.06
0.06-0.07
0.07-0.08
0.08-0.09
0.09-0.10
0.10-0.11
0.11-0.12
0.12-0.13
0.13-0.14
0.14-0.15
0.15-0.16
0.16-0.17
0.17-0.18
0.18-0.19
0.19-0.20
above 0.20
Percent

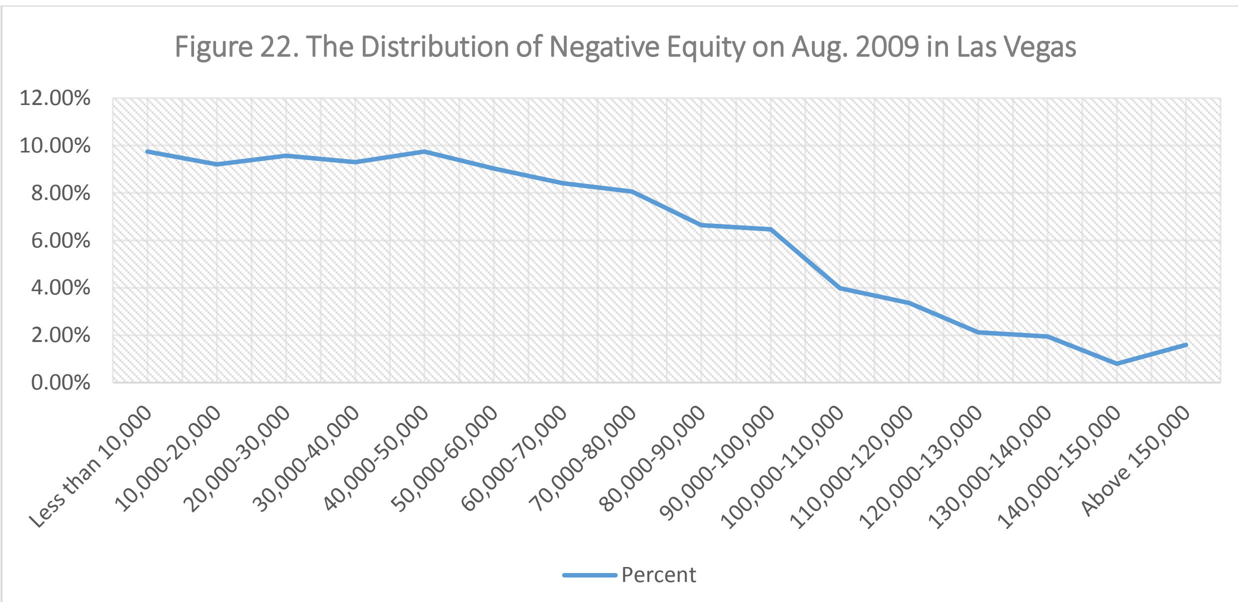


Figure 22. The Distribution of Negative Equity on Aug. 2009 in Las Vegas
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
Less than 10,000
10,000-20,000
20,000-30,000
30,000-40,000
40,000-50,000
50,000-60,000
60,000-70,000
70,000-80,000
80,000-90,000
90,000-100,000
100,000-110,000
110,000-120,000
120,000-130,000
130,000-140,000
140,000-150,000
Above 150,000
Percent

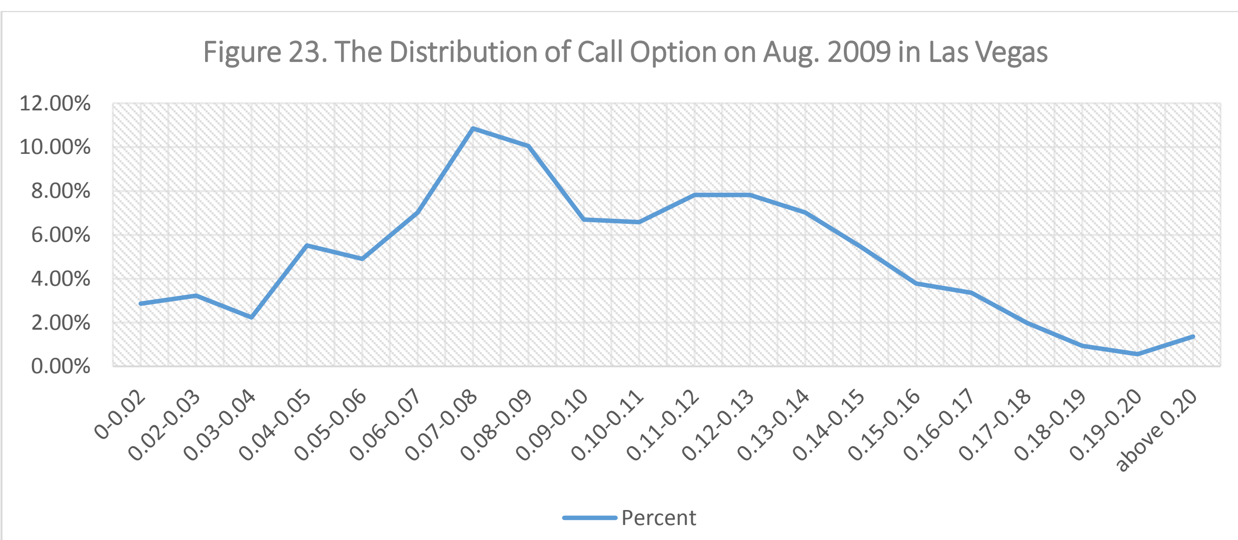


Figure 23. The Distribution of Call Option on Aug. 2009 in Las Vegas
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
0-0.02
0.02-0.03
0.03-0.04
0.04-0.05
0.05-0.06
0.06-0.07
0.07-0.08
0.08-0.09
0.09-0.10
0.10-0.11
0.11-0.12
0.12-0.13
0.13-0.14
0.14-0.15
0.15-0.16
0.16-0.17
0.17-0.18
0.18-0.19
0.19-0.20
above 0.20
Percent

Two different methods are used to deal with this issue. In the first dataset, these 442 mortgages are right censored one month before the modification date. In the second dataset, these 442 mortgages are simply deleted. In the end, the first dataset with right censored mortgages contains a total of 29,741 mortgages; the second dataset without modifications contains 29,299 mortgages. In total, both datasets contain 19,581 mortgages terminated by prepayment, 3,023 mortgages terminated by default and 6,695 mortgages continued without termination[6].

Tables 3 through 6 provide detailed information for the originations and the terminations of mortgages. Table 3 presents the number of mortgages originated per calendar year for the entire dataset (five MSAs); table 4 depicts the number of mortgages originated per calendar year for each MSA; table 5 shows the termination type by calendar year for the entire dataset (five MSAs) and table 6 depicts the termination type by calendar year for each MSA.

[insert Tables 3 through 6 here]

In table 3, findings indicate that about 82 percent of the mortgages in both datasets are originated between 1999 and 2008 and the average number of originations among these years is about 2425 per year. Starting from 2009, the number of originations dramatically decreases at an average rate of 29.7 percent for the right censored dataset and 29.2 percent for the dataset with modified mortgages deleted. The average loan size (base year = 1999) increases from 1999 to 2006 and then decreases from 2006 to 2011.

[6] For termination type, the dataset offers seven different types: prepayment or matured, third party sale prior to 180 days, short sale or short payoff prior to 180 days, deed-in-lieu of foreclosure prior to 180 days, repurchases prior to D180, REO acquisition prior to D180, didn't pay for 180 days. Findings indicate mortgages belong to the type prepayment or matured are all prepayment (without any matured) and the last six types are treated as default. Thus, there are three different termination types at each point in time: prepayment, default and continue.

Table 3. Originations by Calendar Year for the Whole Sample

	#1 dataset with right censored mortgages		#2 dataset with modified mortgages deleted	
Year	Number	Average Loan Size	Number	Average Loan Size
1999	2203	115629.1	2197	115638.6
2000	2423	118768.4	2413	118804.8
2001	2479	126691.4	2472	126749.1
2002	2490	128578	2485	128634.6
2003	2628	132892.6	2613	132959.9
2004	2336	139134.9	2311	138804.7
2005	2794	154240	2743	153661
2006	2509	161082	2392	160286.6
2007	2314	155953.3	2193	155000.9
2008	2293	152549.3	2215	151904.2
2009	1449	150418.7	1443	150276.8
2010	1339	138419.6	1338	138445.9
2011	1511	131707.4	1511	131707.4
2012	718	143266.2	718	143266.2
2013	255	148025.5	255	148025.5

Table 4. Originations by Calendar Year for Each MSA

#1 dataset with right censored mortgages

	AZ Phoenix		FL Miami		FL Tampa		MI Detroit		NV Las Vegas	
Year	No.	Average Loan Size	No.	Average Loan Size	No.	Average Loan Size	No.	Average Loan Size	No.	Average Loan Size
1999	915	119597.8	254	118854.3	388	101567	278	107604.3	368	124423.9
2000	945	125451.9	279	115809.7	488	106097.6	325	112403.2	386	125922.7
2001	1027	129501.3	255	123461.9	460	117196.3	419	122469.8	318	139503.4
2002	895	132955.7	272	132788.9	585	119591.3	376	119756.1	362	138276.1
2003	1029	135329.5	246	149643.4	542	122912	413	123983.2	398	139075.2
2004	803	140274.8	365	151219.1	568	124550.3	228	115775	372	161403.6
2005	1235	161346	333	154437.3	624	136256.5	243	131639.5	359	176167.2
2006	1027	169705.1	288	178436.9	622	145265.1	215	112923.1	357	178835.7
2007	989	165407.4	255	176094.3	520	137094.6	217	110361.5	333	171610.5

2008	1004	157466.1	322	176196.6	415	129070.6	180	112625.6	372	164320.8
2009	777	153364.3	108	163982.8	249	145593.3	76	133537	239	145108.7
2010	711	143328.1	73	158278.6	250	132023.3	62	116981.9	243	130141.8
2011	813	136716	100	165341.6	278	122507.5	82	123859.4	238	113916.2
2012	404	151563	54	171745.6	108	118868.7	41	144701	111	122422.1
2013	160	149931.9	20	181505.9	32	144036.4	12	144699.4	31	121991.3

#2 dataset with modified mortgages deleted

	AZ Phoenix		FL Miami		FL Tampa		MI Detroit		NV Las Vegas	
Year	No.	Average Loan Size	No.	Average Loan Size	No.	Average Loan Size	No.	Average Loan Size	No.	Average Loan Size
1999	913	119584.9	253	118774.7	388	101567	275	107749.1	368	124423.9
2000	940	125544.9	277	115825.1	487	106134.7	324	112427.7	385	125885.5
2001	1026	129493.6	254	123570.3	459	117109.4	415	122797.9	318	139503.4
2002	894	132990.5	272	132788.9	584	119651.8	373	119872.7	362	138276.1
2003	1022	135391.7	245	149747.9	540	123045.2	409	124071.1	397	138983.1
2004	802	140234.2	362	151047	561	124085.9	221	115375.6	365	160330.9
2005	1215	160688	328	154094.5	615	136132.3	237	131495.2	348	174791.2
2006	973	169032.6	276	177997	592	144856.8	208	111924	343	177184.6
2007	936	164526.6	244	176243.4	498	136733	209	110452.2	306	169082.4
2008	974	156713.8	308	176193.7	401	129215.7	173	112191	359	162496.6
2009	774	153160.1	106	164571.2	248	145278.6	76	133537	239	145108.7
2010	710	143384.7	73	158278.6	250	132023.3	62	116981.9	243	130141.8
2011	813	136716	100	165341.6	278	122507.5	82	123859.4	238	113916.2
2012	404	151563	54	171745.6	108	118868.7	41	144701	111	122422.1
2013	160	149931.9	20	181505.9	32	144036.4	12	144699.4	31	121991.3

Table 5. Defaults and Prepayments by Calendar Year for Whole Sample

	Defaults		Prepayments	
Year	Number	Average Loan Size	Number	Average Loan Size
1999	0	0	40	112375
2000	7	109878	217	116721.8
2001	29	122389.8	1650	125209.8

2002	46	101183.2	2077	123492.4
2003	35	105651.3	2935	123770
2004	24	89848.46	1553	119489.5
2005	20	103474.6	1566	127077.8
2006	20	89373.96	1012	129852.5
2007	52	125619.2	883	130372.3
2008	200	151354	650	137238.8
2009	699	157530.9	954	144798.4
2010	833	146844.9	1082	149599
2011	547	139796.4	968	144410.5
2012	407	137479	2664	144001.1
2013	104	134029.4	1330	131012.9

Table 6. Defaults and Prepayments by Calendar Year for Each MSA

AZ Phoenix

	Defaults		Prepayments	
Year	Number	Average Loan Size	Number	Average Loan Size
1999	0	0	14	121928.6
2000	4	127707.3	102	119977
2001	8	122527.7	738	128911.9
2002	13	109561.2	822	127694
2003	15	106660.2	1206	128003.4
2004	9	93584.49	540	121895
2005	3	128241.3	709	132905.6
2006	2	69829.86	467	136315.2
2007	7	126749.6	381	134845.1
2008	64	171915.1	296	147852
2009	248	156626.3	503	152297.6
2010	314	157693.1	540	157700.2
2011	221	148594.9	439	151223.2
2012	151	146629.8	1355	152415.3
2013	32	144148.1	632	132973.3

FL Miami

	Defaults		Prepayments	
Year	Number	Average Loan Size	Number	Average Loan Size

1999	0	0	6	117833.3
2000	1	80300.81	24	116460.4
2001	9	130863.4	127	117507.5
2002	8	109623.5	219	121721.2
2003	2	96428.8	290	122973.6
2004	1	54680.78	211	122381.8
2005	3	164353.7	159	132152.4
2006	2	179739.6	119	131208.3
2007	7	189665.2	125	133955.1
2008	40	192094.2	80	131196.6
2009	141	189507.2	66	157876.2
2010	142	158330.5	82	164479.5
2011	77	144195	89	150251.1
2012	40	142386.5	204	162462.6
2013	10	153114.4	122	143036.6

	FL Tampa			
	Defaults		Prepayments	
Year	Number	Average Loan Size	Number	Average Loan Size
1999	0	0	8	112375
2000	0	0	29	110292.7
2001	2	109592.9	277	114658.1
2002	6	91526.59	394	113140.4
2003	4	115216.6	564	112492.2
2004	3	95544.35	332	107199.2
2005	4	86584.23	334	109376.4
2006	2	101645.8	239	111991.2
2007	15	130997.6	209	114966.5
2008	43	127819.8	138	117257.7
2009	123	141396.3	216	126808.7
2010	125	124896.5	244	134280.3
2011	98	124828.9	230	133242.1
2012	99	115484.9	495	128496.4
2013	29	125176.5	265	121368.3

	MI Detroit			
	Defaults		Prepayments	

Year	Number	Average Loan Size	Number	Average Loan Size
1999	0	0	10	100800
2000	1	41601.63	30	119258
2001	3	103791.8	281	123738.7
2002	14	79642.02	336	120659.2
2003	9	79778.86	397	119713.5
2004	10	90223.29	185	113942.9
2005	9	76868.98	104	112282.3
2006	14	77503.47	50	113926
2007	17	76631.93	62	109907.9
2008	34	84206.89	55	146275.1
2009	74	93270.67	67	131690
2010	90	84237.92	104	139089.1
2011	49	84947.49	95	131421.6
2012	25	69195.95	293	111457.1
2013	9	77316.03	99	110588.8

NV Las Vegas

	Defaults		Prepayments	
Year	Number	Average Loan Size	Number	Average Loan Size
1999	0	0	2	87000
2000	1	136414.6	32	109990.3
2001	7	122964.4	227	132180.3
2002	5	137799.2	306	129913.1
2003	5	145231.7	478	130248.2
2004	1	70555.85	285	130707.8
2005	1	153548.4	260	136740.1
2006	0	0	137	143616.7
2007	6	174931.4	106	152416
2008	19	169745.8	81	132328.2
2009	113	179260.7	102	146061
2010	162	167467.5	112	142777
2011	102	158142.3	115	146950.5
2012	92	162548.7	317	150445.5
2013	24	144550.4	212	139842.9

Table 4 shows the originations for each MSA in both datasets. About 45.5 percent of the total mortgages originate in Phoenix. In 2001, 2003, 2005, 2006 and 2008, over 1,000 new mortgages are originated, and the number of originations gradually decreases starting in 2009. The average loan size in Phoenix reaches $169,705 (base year = 1999) in 2006 and decreases to $136.716 in 2011, then increases again in 2012.

The number of originations in Miami continuously increases from 254 in 1999 to 365 in 2004, and then gradually decreases to 255 in 2007. In 2008 the number of originations in Miami increases again to 322 and then after that year, the number drops below 110. The average loan size in Miami increases from 1999 to 2006, and then continuously decreases until 2010 before it increases again in 2011.

In Tampa, the number of originations gradually increases from 388 in 1999 to 622 in 2006. Starting from 2007, the number of originations in Tampa rapidly decreases at an average rate of 40 percent. The average loan size in Tampa increases from 1999 to 2006 and then decreases from 2007 to 2012, with the exception of 2009, in which loan size increases by 12.8 percent.

The number of originations in Detroit increases from 278 in 1999 to 413 in 2003, then continues to decrease after 2004. Furthermore, in Las Vegas, there are more than 300 originations from 1999 to 2008. The number starts to decrease in 2009.

Table 5 presents aggregate defaults and prepayments for each dataset. The defaults dramatically increase from seven in 2000 to 833 in 2010 and then slightly decrease from 2011 to 2012. The average loan size of defaults peaks in 2009 at around $157,530 and then drops to an average of $139,537 from 2010 to 2013. The prepayments dramatically increase from 40 in 1999 to 2935 in 2003, decrease from 1553 in 2004 to 650 in 2008, and then increase again from 2009 to 2012. The average loan size continued to increase slightly from 1999 to 2001, decreased slightly until 2004, then increased again until 2011.

Table 6 outlines the defaults and prepayments by MSA. The defaults in each MSA are very low from 1999 to 2007 and then increase rapidly from 2008 to 2010. They decrease slightly from 2011. The number of prepayments, from 2001 to around 2005, is large; the number continuously drops from around 2006 until 2011 when it increases again from 2012.

[insert Tables 7 through 8 here]

Tables 7 and 8 discuss the defaults and prepayments by age month for the entire sample and for each individual MSA, respectively. Table 7 indicates a large percentage of the mortgages default between the age month 37 and the age month 48; at the same time, a large amount of mortgages are prepaid between the age month 13 and age month 24.

Table 8 shows us that in Phoenix and Tampa, most mortgages that default do so between the age months 37 and 48. In Miami and Detroit, most mortgages that default do so between the age months 25 and 36 and between the age months 13 and 24, respectively. Furthermore, in Las Vegas, most mortgages that default do so between the age months 49 and 60. In all MSAs, most mortgages that are prepaid end between the age months 13 and 24.

Explanatory Variables, Dependent Variables and Hypotheses

Mortgage information offered by the Freddie Mac dataset includes: the credit score, number of house units, original debt-to-income ratio, original loan-to-value, contract rate, original mortgage term, termination date, termination type, and so on. Not all of this information can be directly used as explanatory variables. As in previous studies, this study uses an indicator

Table 7. Defaults and Prepayments by Age Month for the Whole Sample

	Defaults		Prepayments	
Age Month	Number	Average loan size	Number	Average loan size
1--12	102	177833.3	3451	171847
13--24	459	191100.5	5591	153252
25--36	639	184141	3512	142263.9
37--48	652	187654.2	2247	143191.8
49--60	551	177887.5	1571	146731.1
61--72	379	164556.6	1174	152181.9
73--84	252	144808.1	952	148157.7
85--96	126	131474.5	706	135873
97--`108	54	98523.53	456	120477.7
109--120	26	82647.07	271	109119.2
121--132	15	85938.89	66	84126.88
133--144	6	85825.33	28	77166.42
145--156	3	43944.95	12	74496.54
157--168	1	110295.6	16	76780.59

Table 8. Defaults and Prepayments by Age Month for Each MSA

	AZ Phoenix		FL Miami		FL Tampa		MI Detroit		NV Las Vegas	
Age Month	Default	Prepay ment	Default	Prepay ment	Default	Prepay ment	Default	Prepay ment	Default	Prepay ment
1--12	14	1617	34	251	19	566	28	552	7	465
13--24	156	2349	97	496	66	991	59	554	59	752
25--36	242	1554	116	389	114	728	51	319	116	522
37--48	287	1027	88	244	123	451	49	183	105	342
49--60	215	719	73	150	93	338	51	127	119	237
61--72	139	499	50	130	81	293	38	115	71	137
73--84	73	419	28	85	59	246	43	87	49	115
85--96	36	292	13	87	24	165	28	72	25	90
97--`108	11	148	5	51	12	112	17	77	9	68
109--120	3	91	1	27	6	53	10	69	6	31
121--132	1	19	1	9	3	22	7	9	3	7
133--144	0	11	3	3	1	6	0	4	2	2
145--156	2	4	0	1	0	3	0	1	1	3
157--168	1	5	0	3	0	3	0	2	0	3

for a prepayment penalty and the value of the call option to explain the prepayment risk. The unemployment rate, negative equity, a dummy for negative equity and original loan-to-value are used to explain the default risk. Other explanatory variables for the prepayment risk and the default risk are the debt-to-income ratio and the credit score, log mortgage size, current loan age in months, a dummy for house unit and time period dummies. Table 9 lists the summary statistics for important explanatory variables in each MSA. Tables 10 through 14 summarize the definition of explanatory variables in each MSA.

[insert Table 9 through 14 here]

The indicator for a prepayment penalty equals to one when there is an additional charge for borrowers to prepay their mortgages. Otherwise, the indicator equals to zero. Therefore, this a research posits the following hypothesis:

H1: There is a negative relationship between the prepayment penalty indicator and the prepayment risk.

The value of the call option is the difference between the market value of the mortgage and the book value of the mortgage measured as a percent of the market value of the mortgage. The calculation is presented in the Appendix A. The value of the call option is "in the money," when the market value of the mortgage is larger than its book value. In this situation, borrowers are more likely to prepay their current mortgage and refinance because the value of the new mortgages against the housing equity would increase. Therefore, hypothesis 2 posits:

H2: A positive relationship exists between the value of the call option and the prepayment risk.

Table 9. Summary Statistics for Each MSA

Phoenix: Prepayment and default are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right censored mortgages	Call option	494949	0.0396204	0.0858311	-0.2811943	0.3463946
	The unemployment rate	494949	6.213542	2.397017	2.6	10.5
	Negative equity	494949	8.979612	22.57022	0	201.6359
	Original loan-to-value	12734	0.7350212	0.1562165	0.08	1
	Log loan size	12734	11.9329	0.5073528	9.472705	12.94084
	Loan age	494949	31.01697	26.23423	1	169
	The credit score	12699	0.7298035	0.0553845	0.496	0.84
	The debt-to-income ratio	12508	0.341563	0.1189343	0	0.65
#2 dataset with modified mortgages deleted	Call option	487503	.0396846	.0858345	-.2813612	.3463946
	The unemployment rate	487503	6.213595	2.39288	2.6	10.5
	Negative equity	487503	8.766227	22.24506	0	201.6359
	Original loan-to-value	12556	0.734686	0.156618	0.08	1
	Log loan size	12556	11.92892	0.50758	9.472705	12.94084
	Loan age	487503	31.10495	26.32927	1	169
	The credit score	12522	0.730355	0.055221	0.496	0.84
	The debt-to-income ratio	12331	0.340491	0.118921	0	0.65
Phoenix: Prepayment and 90-days-delinquency are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right censored mortgages	Call option	486546	0.038343	0.085356	-0.28119	0.346395
	The unemployment rate	486546	6.188801	2.392581	2.6	10.5
	Negative equity	486546	8.635514	22.16521	0	201.6359
	Original loan-to-value	12698	0.734854	0.156323	0.08	1
	Log loan size	12698	11.93217	0.507398	9.472705	12.94084
	Loan age	486546	30.73576	26.12315	1	169
	The credit score	12663	0.729842	0.055362	0.496	0.84
	The debt-to-income ratio	12472	0.341376	0.118877	0	0.65
#2 dataset with modified mortgages deleted	Call option	479523	0.038459	0.085374	-0.28136	0.346395
	The unemployment rate	479523	6.190508	2.388777	2.6	10.5
	Negative equity	479523	8.435202	21.84729	0	201.6359
	Original loan-to-value	12522	0.734532	0.15672	0.08	1
	Log loan size	12522	11.92822	0.507627	9.472705	12.94084
	Loan age	479523	30.8299	26.21692	1	169
	The credit score	12488	0.730387	0.055197	0.496	0.84
	The debt-to-income ratio	12297	0.340302	0.118854	0	0.65
Miami: Prepayment and default are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right	Call option	150698	0.049442	0.088105	-0.2328	0.46051

censored mortgages	The unemployment rate	150698	6.943224	2.77443	3.3	11.9
	Negative equity	150698	6.097388	18.07763	0	187.1141
	Original loan-to-value	3224	0.741331	0.175259	0.08	1
	Log loan size	3224	11.95617	0.514616	9.903487	13.02805
	Loan age	150698	35.7997	29.23231	1	169
	The credit score	3184	0.70726	0.055179	0.496	0.817
	The debt-to-income ratio	3155	0.349794	0.12597	0	0.65
#2 dataset with modified mortgages deleted	Call option	148637	0.049518	0.088076	-0.23296	0.46051
	The unemployment rate	148637	6.940274	2.771599	3.3	11.9
	Negative equity	148637	6.057282	18.00889	0	187.1141
	Original loan-to-value	3172	0.74134	0.175783	0.08	1
	Log loan size	3172	11.95256	0.51559	9.903487	13.02805
	Loan age	148637	35.90544	29.29226	1	169
	The credit score	3132	0.70795	0.055132	0.496	0.817
	The debt-to-income ratio	3104	0.348853	0.126003	0	0.65
Miami: Prepayment and 90-days-delinquency are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right censored mortgages	Call option	147098	0.048015	0.087727	-0.2328	0.46051
	The unemployment rate	147098	6.891245	2.761292	3.3	11.9
	Negative equity	147098	5.755085	17.53271	0	167.8424
	Original loan-to-value	3211	0.740987	0.175413	0.08	1
	Log loan size	3211	11.95622	0.51491	9.903487	13.02805
	Loan age	147098	35.55935	29.21223	1	169
	The credit score	3171	0.707351	0.055221	0.496	0.817
	The debt-to-income ratio	3143	0.349634	0.125679	0	0.65
#2 dataset with modified mortgages deleted	Call option	145222	0.048162	0.087699	-0.23296	0.46051
	The unemployment rate	145222	6.890992	2.759121	3.3	11.9
	Negative equity	145222	5.722438	17.48247	0	167.8424
	Original loan-to-value	3160	0.741057	0.175953	0.08	1
	Log loan size	3160	11.95263	0.515812	9.903487	13.02805
	Loan age	145222	35.67159	29.26933	1	169
	The credit score	3120	0.708031	0.05517	0.496	0.817
	The debt-to-income ratio	3093	0.348652	0.125701	0	0.65
Tampa: Prepayment and default are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right censored mortgages	Call option	283421	0.045595	0.088259	-0.24649	0.328896
	The unemployment rate	283421	7.101306	3.128368	2.6	12.5
	Negative equity	283421	4.658355	13.12679	0	134.6386
	Original loan-to-value	6129	0.744368	0.153808	0.13	1
	Log loan size	6129	11.76994	0.529389	9.472705	12.94084
	Loan age	283421	34.73948	27.69508	1	169
	The credit score	6105	0.720715	0.057229	0.453	0.835

	The debt-to-income ratio	5991	0.346508	0.121579	0	0.65
#2 dataset with modified mortgages deleted	Call option	279662	0.045752	0.088327	-0.24666	0.328896
	The unemployment rate	279662	7.098834	3.127134	2.6	12.5
	Negative equity	279662	4.60253	13.06438	0	134.6386
	Original loan-to-value	6041	0.743753	0.154406	0.13	1
	Log loan size	6041	11.76655	0.53047	9.472705	12.94084
	Loan age	279662	34.85329	27.77742	1	169
	The credit score	6017	0.721421	0.056994	0.482	0.835
	The debt-to-income ratio	5903	0.345092	0.121243	0	0.65
Tampa: Prepayment and 90-days-delinquency are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right censored mortgages	Call option	278744	0.044472	0.087916	-0.24649	0.328896
	The unemployment rate	278744	7.06956	3.12483	2.6	12.5
	Negative equity	278744	4.536016	12.97167	0	134.6386
	Original loan-to-value	6112	0.744195	0.153928	0.13	1
	Log loan size	6112	11.76895	0.529016	9.472705	12.94084
	Loan age	278744	34.47394	27.64089	1	169
	The credit score	6088	0.720737	0.057233	0.453	0.835
	The debt-to-income ratio	5975	0.346474	0.121553	0	0.65
#2 dataset with modified mortgages deleted	Call option	275153	0.044653	0.087985	-0.24666	0.328896
	The unemployment rate	275153	7.068366	3.123925	2.6	12.5
	Negative equity	275153	4.482206	12.90979	0	134.6386
	Original loan-to-value	6024	0.743576	0.154528	0.13	1
	Log loan size	6024	11.76554	0.530089	9.472705	12.94084
	Loan age	275153	34.59042	27.72149	1	169
	The credit score	6000	0.721445	0.056998	0.482	0.835
	The debt-to-income ratio	5887	0.345054	0.121215	0	0.65
Detroit: Prepayment and default are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right censored mortgages	Call option	144368	0.04301	0.090517	-0.2796	0.367435
	The unemployment rate	144368	9.12082	3.289267	3	16.9
	Negative equity	144368	5.322262	12.39206	0	134.7806
	Original loan-to-value	3165	0.741652	0.14927	0.14	1
	Log loan size	3167	11.68306	0.517123	9.305651	12.94084
	Loan age	144368	38.32582	30.78718	1	169
	The credit score	3152	0.717175	0.058886	0.407	0.823
	The debt-to-income ratio	3108	0.334569	0.118309	0.01	0.65
#2 dataset with modified mortgages deleted	Call option	141824	0.043238	0.090613	-0.27977	0.367435
	The unemployment rate	141824	9.122959	3.291029	3	16.9
	Negative equity	141824	5.272115	12.27291	0	134.7806
	Original loan-to-value	3115	0.740562	0.149689	0.14	1

	Log loan size	3117	11.68194	0.518284	9.305651	12.94084
	Loan age	141824	38.40756	30.85015	1	169
	The credit score	3102	0.717755	0.058541	0.493	0.823
	The debt-to-income ratio	3060	0.333967	0.118285	0.01	0.65
Detroit: Prepayment and 90-days-delinquency are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right censored mortgages	Call option	140309	0.041293	0.089971	-0.2796	0.367435
	The unemployment rate	140309	9.085858	3.281992	3	16.9
	Negative equity	140309	5.216362	12.29911	0	134.7806
	Original loan-to-value	3150	0.741495	0.149455	0.14	1
	Log loan size	3152	11.68396	0.516829	9.305651	12.94084
	Loan age	140309	37.79644	30.58658	1	169
	The credit score	3137	0.717303	0.058849	0.407	0.823
	The debt-to-income ratio	3093	0.334265	0.118165	0.01	0.65
#2 dataset with modified mortgages deleted	Call option	138144	0.04152	0.090033	-0.27977	0.367435
	The unemployment rate	138144	9.088978	3.283736	3	16.9
	Negative equity	138144	5.164517	12.17531	0	134.7806
	Original loan-to-value	3101	0.740416	0.149857	0.14	1
	Log loan size	3103	11.68313	0.51814	9.305651	12.94084
	Loan age	138144	37.93369	30.6864	1	169
	The credit score	3088	0.717864	0.058506	0.493	0.823
	The debt-to-income ratio	3046	0.333667	0.118119	0.01	0.65
Las Vegas: Prepayment and default are dependent variables						
Dataset	Explanatory variables	No. of observations	Mean	St. Dev.	Min	Max
#1 dataset with right censored mortgages	Call option	188697	0.041867	0.086761	-0.27907	0.327644
	The unemployment rate	188697	8.439248	3.889642	3.8	14.6
	Negative equity	188697	13.73526	29.78763	0	209.7891
	Original loan-to-value	4487	0.748908	0.153044	0.09	1
	Log loan size	4487	11.95056	0.478449	9.392662	12.94084
	Loan age	188697	32.49552	27.02609	1	170
	The credit score	4472	0.725283	0.05558	0.514	0.823
	The debt-to-income ratio	4398	0.35721	0.120773	0	0.65
#2 dataset with modified mortgages deleted	Call option	185452	0.042186	0.086851	-0.27923	0.327644
	The unemployment rate	185452	8.438328	3.891042	3.8	14.6
	Negative equity	185452	13.36946	29.33581	0	209.7891
	Original loan-to-value	4413	0.74845	0.153573	0.09	1
	Log loan size	4413	11.94304	0.476878	9.392662	12.94084
	Loan age	185452	32.60197	27.13556	1	170
	The credit score	4398	0.725793	0.055546	0.514	0.823
	The debt-to-income ratio	4324	0.356568	0.120729	0	0.65
Las Vegas: Prepayment and 90-days-delinquency are dependent variables						
Dataset	Explanatory variables	No. of	Mean	St. Dev.	Min	Max

		observations				
#1 dataset with right censored **mortgages**	Call option	185490	0.040767	0.086501	-0.27907	0.327644
	The unemployment rate	185490	8.395085	3.881563	3.8	14.6
	Negative equity	185490	13.20461	29.24027	0	209.7891
	Original loan-to-value	4473	0.748876	0.15304	0.09	1
	Log loan size	4473	11.94997	0.478825	9.392662	12.94084
	Loan age	185490	32.26378	26.98676	1	170
	The credit score	4458	0.725387	0.055541	0.514	0.823
	The debt-to-income ratio	4384	0.357224	0.120768	0	0.65
#2 dataset with modified mortgages deleted	Call option	182423	0.041106	0.086585	-0.27923	0.327644
	The unemployment rate	182423	8.39682	3.883443	3.8	14.6
	Negative equity	182423	12.86818	28.80553	0	209.7891
	Original loan-to-value	4402	0.748485	0.153552	0.09	1
	Log loan size	4402	11.94266	0.477281	9.392662	12.94084
	Loan age	182423	32.36745	27.09506	1	170
	The credit score	4387	0.725886	0.055514	0.514	0.823
	The debt-to-income ratio	4313	0.356587	0.120722	0	0.65

Table 10. Explanatory Variables Definition for Phoenix

Variable Name	Variable Definition for Different Dependent Variables	
	Prepayment and default	Prepayment and 90-days-delinquency
Prepayment penalty	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.
Call option variables		
Call option part 1	Call option value below -0.0247799 for dataset which is right censored and Call option value below -0.0246538 for dataset in which modification mortgages are deleted (-0.0247799 and -0.0246538 are 25th percentiles of Call option value in two dataset respectively)	Call option value below -0.0252991 for dataset which is right censored and Call option value below -0.0251161 for dataset in which modification mortgages are deleted (-0.0252991 and -0.0251161 are 25th percentiles of Call option value in two dataset respectively)
Call option part 2	Call option value is from -0.0247799 to 0.0365862 for dataset which is right censored and Call option value is from -0.0246538 to 0.0367869 for dataset in which modification mortgages are deleted (0.0365862 and 0.0367869 are 50th percentiles of Call option value in two dataset respectively)	Call option value is from -0.0252991 to 0.0353509 for dataset which is right censored and Call option value is from -0.0251161 to 0.0355266 for dataset in which modification mortgages are deleted (0.0353509 and 0.0355266 are 50th percentiles of Call option value in two dataset respectively)
Call option part 3	Call option value is from 0.0365862 to 0.1020157 for dataset which is right censored and Call option value is 0.0367869 to 0.1020334 for dataset in which modification mortgages are deleted (0.1020157 and 0.1020334 are 75th percentiles of Call option value in two dataset respectively)	Call option value is from 0.0353509 to 0.1003149 for dataset which is right censored and Call option value is 0.0355266 to 0.1003559 for dataset in which modification mortgages are deleted (0.1003149 and 0.1003559 are 75th percentiles of Call option value in two dataset respectively)
Call option part 4	Call option value above 0.1020157 for dataset	Call option value above 0.1003149 for dataset

	which is right censored and Call option value above 0.1020334 for dataset in which modification mortgages are deleted	which is right censored and Call option value above 0.1003559 for dataset in which modification mortgages are deleted
The unemployment rate variables		
The unemployment rate part 1	The unemployment rate in Phoenix below 5.7 (5.7 is 50th percentile of the unemployment rate in both datasets)	The unemployment rate in Phoenix below 5.7 (5.7 is 50th percentile of the unemployment rate in both datasets)
The unemployment rate part 2	The unemployment rate in Phoenix above 5.7	The unemployment rate in Phoenix above 5.7
Negative equity variables		
Negative equity part 1	Negative equity value below 18.36619 for dataset which is right censored and Negative equity value below 18.25616 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 18.36619 and 18.25616 are 25th percentiles of negative equity value in two dataset respectively)	Negative equity value below 18.02891 for dataset which is right censored and Negative equity value below 17.93523 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 18.02891 and 17.93523 are 25th percentiles of negative equity value in two dataset respectively)
Negative equity part 2	Negative equity value is from 18.36619 to 37.80385 for dataset which is right censored and Negative equity value is from 18.25616 to 37.61306 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 37.80385 and 37.61306 are 50th percentiles of negative equity value in two dataset respectively)	Negative equity value is from 18.02891 to 37.33448 for dataset which is right censored and Negative equity value is from 17.93523 to 37.15196 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 37.33448 and 37.15196 are 50th percentiles of negative equity value in two dataset respectively)
Negative equity part 3	Negative equity value is from 37.80385 to 62.90972 for dataset which is right censored and Negative equity value is from 37.61306 to 62.59928 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 62.90972 and 62.59928 are 75th percentiles of negative equity value in two dataset respectively)	Negative equity value is from 37.33448 to 62.50297 for dataset which is right censored and Negative equity value is from 37.15196 to 62.15913 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 62.50297 and 62.15913 are 75th percentiles of negative equity value in two dataset respectively)
Negative equity part 4	Negative equity value above 62.90972 for dataset which is right censored and Negative equity value above 62.59928 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)	Negative equity value above 62.50297 for dataset which is right censored and Negative equity value above 62.15913 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)
The negative equity dummy	Equal to 1 if the equity less than zero in that month and 0 otherwise.	Equal to 1 if the equity less than zero in that month and 0 otherwise.
Original loan-to-value variables		
Original LTV part 1	Original loan-to-value below 0.65 (Original loan-to-value is divided by 100; 0.65 is 25th percentiles of original loan-to-value in both datasets)	Original loan-to-value below 0.65 (Original loan-to-value is divided by 100; 0.65 is 25th percentiles of original loan-to-value in both datasets)
Original LTV part 2	Original loan-to-value is from 0.65 to 0.78 (0.78 is 50th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.65 to 0.78 (0.78 is 50th percentiles of original loan-to-value in both datasets)
Original LTV part 3	Original loan-to-value is from 0.78 to 0.80 (0.80 is 75th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.78 to 0.80 (0.80 is 75th percentiles of original loan-to-value in both datasets)
Original LTV part 4	Original loan-to-value is from 0.80 to 0.95 (0.95 is 95th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.80 to 0.95 (0.95 is 95th percentiles of original loan-to-value in both datasets)
Original LTV part 5	Original loan-to-value above 0.95	Original loan-to-value above 0.95

Log loan size variables		
Log loan size part 1	Log loan size value is taking log of loan size. Log loan size value below 11.56172 (11.56172 is 25th percentiles of log loan size value in both datasets)	Log loan size value is taking log of loan size. Log loan size value below 11.56172 for dataset which is right censored and Log loan size value below 11.55215 for dataset in which modification mortgages are deleted (11.56172 and 11.55215 are 25th percentiles of log loan size value in both datasets)
Log loan size part 2	Log loan size value is from 11.56172 to 11.91839 (11.91839 is 50th percentiles of log loan size value in both datasets)	Log loan size value is from 11.56172 to 11.91839 for dataset which is right censored and from 11.55215 to 11.91839 for dataset in which modification mortgages are deleted (11.91839 is 50th percentiles of log loan size value in both datasets)
Log loan size part 3	Log loan size value is from 11.91839 to 12.25486 for dataset which is right censored and to 12.24529 for dataset in which modification mortgages are deleted (12.25486 and 12.24529 are 75th percentiles of log loan size value in two dataset respectively)	Log loan size value is from 11.91839 to 12.25009 for dataset which is right censored and from 11.91839 to 12.24529 for dataset in which modification mortgages are deleted (12.25009 and 12.24529 are 75th percentiles of log loan size value in two dataset respectively)
Log loan size part 4	Log loan size value above 12.25486 for dataset which is right censored and Log loan size value above 12.24529 for dataset in which modification mortgages are deleted	Log loan size value above 12.25009 for dataset which is right censored and Log loan size value above 12.24529 for dataset in which modification mortgages are deleted
Loan age variables		
Loan age part 1	Loan age value below 11 (11 is 25th percentiles of loan age value in both datasets)	Loan age value below 11 (11 is 25th percentiles of loan age value in both datasets)
Loan age part 2	Loan age value is from 11 to 24 (24 is 50th percentiles of loan age value in both datasets)	Loan age value is from 11 to 23 (23 is 50th percentiles of loan age value in both datasets)
Loan age part 3	Loan age value is from 24 to 45 (45 is 75th percentiles of loan age value in both datasets)	Loan age value is from 23 to 44 (44 is 75th percentiles of loan age value in both datasets)
Loan age part 4	Loan age value above 45	Loan age value above 44
The credit score variables		
The credit score part 1	The credit score value below 0.688 (The credit score value is divided by 1000; 0.688 is 25th percentiles of the credit score value in both datasets)	The credit score value below 0.689 for dataset which is right censored and The credit score value below 0.69 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.689 and 0.69 are 25th percentiles of the credit score value in two dataset respectively)
The credit score part 2	The credit score value is from 0.688 to 0.736 for dataset which is right censored and to 0.737 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.736 and 0.737 are 50th percentiles of the credit score value in two dataset respectively)	The credit score value is from 0.689 to 0.737 for dataset which is right censored and from 0.69 to 0.738 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.737 and 0.738 are 50th percentiles of the credit score value in two dataset respectively)
The credit score part 3	The credit score value is from 0.736 to 0.774 for dataset which is right censored and from 0.737 to 0.774 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.774 is 75th percentiles of the credit score value in both datasets)	The credit score value is from 0.737 to 0.774 for dataset which is right censored and from 0.738 to 0.774 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.774 is 75th percentiles of the credit score value in both datasets)
The credit score part 4	The credit score value above 0.774 (The credit score value is divided by 1000)	The credit score value above 0.774 (The credit score value is divided by 1000)
The debt-to-income ratio		

variables		
DTI part 1	The debt-to-income ratio value below 0.25 (The debt-to-income ratio value is divided by 100; 0.25 is 25th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value below 0.25 (The debt-to-income ratio value is divided by 100; 0.25 is 25th percentiles of the debt-to-income ratio value in both datasets)
DTI part 2	The debt-to-income ratio value is from 0.25 to 0.34 (The debt-to-income ratio value is divided by 100; 0.34 is 50th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.25 to 0.34 for dataset which is right censored and to 0.33 for dataset in which modification mortgages are deleted (The debt-to-income ratio value is divided by 100; 0.34 and 0.33 are 50th percentiles of the debt-to-income ratio value in two dataset respectively)
DTI part 3	The debt-to-income ratio value is from 0.34 to 0.42 (The debt-to-income ratio value is divided by 100; 0.42 is 75th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.34 to 0.42 for dataset which is right censored and from 0.33 to 0.42 for dataset in which modification mortgages are deleted (The debt-to-income ratio value is divided by 100; 0.42 is 75th percentiles of the debt-to-income ratio value in both datasets)
DTI part 4	The debt-to-income ratio value above 0.42 (The debt-to-income ratio value is divided by 100)	The debt-to-income ratio value above 0.42 (The debt-to-income ratio value is divided by 100)
The dummy for the number of units	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1
Time period variables		
Time period part 1	Time period starts from March 1999 to July 2003 for prepayment and time period start from March 1999 to January 2008 for default	Time period starts from March 1999 to July 2003 for prepayment and time period start from March 1999 to January 2008 for default
Time period part 2	Time period starts from July 2003 to November 2008 for prepayment and time period start from January 2008 to June 2010 for default	Time period starts from July 2003 to November 2008 for prepayment and time period start from January 2008 to June 2010 for default
Time period part 3	Time period starts from November 2008 to March 2013 for prepayment and time period start from June 2010 to March 2013 for default	Time period starts from November 2008 to March 2013 for prepayment and time period start from June 2010 to March 2013 for default

Table 11. Explanatory Variables Definition for Miami

Variable Name	Variable Definition for Different Dependent Variables	
	Prepayment and default	Prepayment and 90-days-delinquency
Prepayment penalty	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.
Call option variables		
Call option part 1	Call option value below -0.0164943 for dataset which is right censored and Call option value below -0.0160465 for dataset in which modification mortgages are deleted (-0.0164943 and -0.0160465 are 25th percentiles of Call option value in two dataset respectively)	Call option value below -0.0176692 for dataset which is right censored and Call option value below -0.0174106 for dataset in which modification mortgages are deleted (-0.0176692 and -0.0174106 are 25th percentiles of Call option value in two dataset respectively)
Call option part 2	Call option value is from -0.0164943 to 0.0484381 for dataset which is right censored and Call option value is from -0.0160465 to 0.0486717 for dataset in which modification mortgages are deleted	Call option value is from -0.0176692 to 0.0466797 for dataset which is right censored and Call option value is from -0.0174106 to 0.0469463 for dataset in which modification mortgages are deleted

	(0.0484381 and 0.0486717 are 50th percentiles of Call option value in two dataset respectively)	(0.0466797 and 0.0469463 are 50th percentiles of Call option value in two dataset respectively)
Call option part 3	Call option value is from 0.0484381 to 0.1149434 for dataset which is right censored and Call option value is from 0.0486717 to 0.1150138 for dataset in which modification mortgages are deleted (0.1149434 and 0.1150138 are 75th percentiles of Call option value in two dataset respectively)	Call option value is from 0.0466797 to 0.1131345 for dataset which is right censored and Call option value is from 0.0469463 to 0.113277 for dataset in which modification mortgages are deleted (0.1131345 and 0.113277 are 75th percentiles of Call option value in two dataset respectively)
Call option part 4	Call option value above 0.1149434 for dataset which is right censored and Call option value above 0.1150138 for dataset in which modification mortgages are deleted	Call option value above 0.1131345 for dataset which is right censored and Call option value above 0.113277 for dataset in which modification mortgages are deleted
The unemployment rate variables		
The unemployment rate part 1	The unemployment rate in Miami below 6.2 for dataset which is right censored and below 6.1 for dataset in which modification mortgages are deleted (6.2 and 6.1 are 50th percentile of the unemployment rate in two dataset respectively)	The unemployment rate in Miami below 6.1 (6.1 is 50th percentile of the unemployment rate in both datasets)
The unemployment rate part 2	The unemployment rate in Miami above 6.2 for dataset which is right censored and above 6.1 for dataset in which modification mortgages are deleted	The unemployment rate in Miami above 6.1
Negative equity variables		
Negative equity part 1	Negative equity value below 13.19394 for dataset which is right censored and Negative equity value below 13.18863 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 13.19394 and 13.18863 are 25th percentiles of negative equity value in two dataset respectively)	Negative equity value below 12.90738 for dataset which is right censored and Negative equity value below 12.92502 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 12.90738 and 12.92502 are 25th percentiles of negative equity value in two dataset respectively)
Negative equity part 2	Negative equity value is from 13.19394 to 30.56755 for dataset which is right censored and Negative equity value is from 13.18863 to 30.44288 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 30.56755 and 30.44288 are 50th percentiles of negative equity value in two dataset respectively)	Negative equity value is from 12.90738 to 29.92025 for dataset which is right censored and Negative equity value is from 12.92502 to 29.80616 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 29.92025 and 29.80616 are 50th percentiles of negative equity value in two dataset respectively)
Negative equity part 3	Negative equity value is from 30.56755 to 53.85436 for dataset which is right censored and Negative equity value is from 30.44288 to 53.75649 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 53.85436 and 53.75649 are 75th percentiles of negative equity value in two dataset respectively)	Negative equity value is from 29.92025 to 53.0423 for dataset which is right censored and Negative equity value is from 29.80616 to 53.07203 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 53.0423 and 53.07203 are 75th percentiles of negative equity value in two dataset respectively)
Negative equity part 4	Negative equity value above 53.85436 for dataset which is right censored and Negative equity value above 53.75649 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)	Negative equity value above 53.0423 for dataset which is right censored and Negative equity value above 53.07203 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)

The negative equity dummy	Equal to 1 if the equity less than zero in that month and 0 otherwise.	Equal to 1 if the equity less than zero in that month and 0 otherwise.
Original loan-to-value variables		
Original LTV part 1	Original loan-to-value below 0.63 (Original loan-to-value is divided by 100; 0.63 is 25th percentiles of original loan-to-value in both datasets)	Original loan-to-value below 0.63 (Original loan-to-value is divided by 100; 0.63 is 25th percentiles of original loan-to-value in both datasets)
Original LTV part 2	Original loan-to-value is from 0.63 to 0.77 (0.77 is 50th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.63 to 0.77 (0.77 is 50th percentiles of original loan-to-value in both datasets)
Original LTV part 3	Original loan-to-value is from 0.77 to 0.80 (0.80 is 75th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.77 to 0.80 (0.80 is 75th percentiles of original loan-to-value in both datasets)
Original LTV part 4	Original loan-to-value is from 0.80 to 0.95 (0.95 is 95th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.80 to 0.95 (0.95 is 95th percentiles of original loan-to-value in both datasets)
Original LTV part 5	Original loan-to-value above 0.95	Original loan-to-value above 0.95
Log loan size variables		
Log loan size part 1	Log loan size value is taking log of loan size. Log loan size value below 11.60824 for dataset which is right censored and Log loan size value below 11.5991 for dataset in which modification mortgages are deleted (11.60824 and 11.5991 are 25th percentiles of log loan size value in two dataset respectively)	Log loan size value is taking log of loan size. Log loan size value below 11.5991 for dataset which is right censored and Log loan size value below 11.58989 for dataset in which modification mortgages are deleted (11.5991 and 11.58989 are 25th percentiles of log loan size value in both datasets)
Log loan size part 2	Log loan size value is from 11.60824 to 11.95118 for dataset which is right censored and Log loan size value from 11.5991 to 11.95118 for dataset in which modification mortgages are deleted (11.95118 is 50th percentiles of log loan size value in both datasets)	Log loan size value is from 11.5991 to 11.94471 for dataset which is right censored and from 11.58989 to 11.94471 for dataset in which modification mortgages are deleted (11.94471 is 50th percentiles of log loan size value in both datasets)
Log loan size part 3	Log loan size value is from 11.95118 to 12.30138 (12.30138 is 75th percentiles of log loan size value in both datasets)	Log loan size value is from 11.94471 to 12.29683 for dataset which is right censored and from 11.94471 to 12.28765 for dataset in which modification mortgages are deleted (12.29683 and 12.28765 are 75th percentiles of log loan size value in two dataset respectively)
Log loan size part 4	Log loan size value above 12.30138	Log loan size value above 12.29683 for dataset which is right censored and Log loan size value above 12.28765 for dataset in which modification mortgages are deleted
Loan age variables		
Loan age part 1	Loan age value below 13 (13 is 25th percentiles of loan age value in both datasets)	Loan age value below 12 for dataset which is right censored and below 13 for dataset in which modification mortgages are deleted (12 and 13 are 25th percentiles of loan age value in two dataset respectively)
Loan age part 2	Loan age value is from 13 to 28 (28 is 50th percentiles of loan age value in both datasets)	Loan age value from 12 to 28 for dataset which is right censored and from 13 to 28 for dataset in which modification mortgages are deleted (28 is 50th percentiles of loan age value in both datasets)
Loan age part 3	Loan age value is from 28 to 52 for dataset which is right censored and to 53 for dataset in which modification mortgages are deleted (52 and 53 are 75th percentiles of loan age value in two dataset respectively)	Loan age value is from 28 to 52 (52 is 75th percentiles of loan age value in both datasets)
Loan age part 4	Loan age value above 52 for dataset which is	Loan age value above 52

	right censored and above 53 for dataset in which modification mortgages are deleted	
The credit score variables		
The credit score part 1	The credit score value below 0.671 for dataset which is right censored and below 0.672 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.671 and 0.672 are 25th percentiles of the credit score value in two dataset respectively)	The credit score value below 0.671 for dataset which is right censored and The credit score value below 0.672 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.671 and 0.672 are 25th percentiles of the credit score value in two dataset respectively)
The credit score part 2	The credit score value is from 0.671 to 0.71 for dataset which is right censored and from 0.672 to 0.711 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.71 and 0.711 are 50th percentiles of the credit score value in two dataset respectively)	The credit score value is from 0.671 to 0.711 for dataset which is right censored and from 0.672 to 0.712 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.711 and 0.712 are 50th percentiles of the credit score value in two dataset respectively)
The credit score part 3	The credit score value is from 0.71 to 0.753 for dataset which is right censored and from 0.711 to 0.754 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.753 and 0.754 are 75th percentiles of the credit score value in two dataset respectively)	The credit score value is from 0.711 to 0.753 for dataset which is right censored and from 0.712 to 0.754 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.753 and 0.754 are 75th percentiles of the credit score value in two dataset respectively)
The credit score part 4	The credit score value above 0.753 for dataset which is right censored and 0.754 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000)	The credit score value above 0.753 for dataset which is right censored and above 0.754 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000)
The debt-to-income ratio variables		
DTI part 1	The debt-to-income ratio value below 0.25 (The debt-to-income ratio value is divided by 100; 0.25 is 25th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value below 0.25 (The debt-to-income ratio value is divided by 100; 0.25 is 25th percentiles of the debt-to-income ratio value in both datasets)
DTI part 2	The debt-to-income ratio value is from 0.25 to 0.35 for dataset which is right censored and to 0.34 for dataset in which modification mortgages are deleted (The debt-to-income ratio value is divided by 100; 0.35 and 0.34 are 50th percentiles of the debt-to-income ratio value in two dataset respectively)	The debt-to-income ratio value is from 0.25 to 0.34 (The debt-to-income ratio value is divided by 100; 0.34 is 50th percentiles of the debt-to-income ratio value in both datasets)
DTI part 3	The debt-to-income ratio value is from 0.35 to 0.43 for dataset which is right censored and from 0.34 to 0.43 for dataset in which modification mortgages are deleted (The debt-to-income ratio value is divided by 100; 0.43 is 75th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.34 to 0.43 (The debt-to-income ratio value is divided by 100; 0.43 is 75th percentiles of the debt-to-income ratio value in both datasets)
DTI part 4	The debt-to-income ratio value above 0.43 (The debt-to-income ratio value is divided by 100)	The debt-to-income ratio value above 0.43 (The debt-to-income ratio value is divided by 100)
The dummy for the number of units	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1
Time period variables		
Time period part 1	Time period starts from March 1999 to April 2003 for prepayment and time period start from March 1999 to July 2008 for default	Time period starts from March 1999 to April 2003 for prepayment and time period start from March 1999 to July 2008 for default

Time period part 2	Time period starts from April 2003 to December 2008 for prepayment and time period start from July 2008 to February 2010 for default	Time period starts from April 2003 to December 2008 for prepayment and time period start from July 2008 to February 2010 for default
Time period part 3	Time period starts from December 2008 to March 2013 for prepayment and time period start from February 2010 to March 2013 for default	Time period starts from December 2008 to March 2013 for prepayment and time period start from February 2010 to March 2013 for default

Table 12. Explanatory Variables Definition for Tampa

Variable Name	Variable Definition for Different Dependent Variables	
	Prepayment and default	Prepayment and 90-days-delinquency
Prepayment penalty	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.
Call option variables		
Call option part 1	Call option value below -0.0206478 for dataset which is right censored and Call option value below -0.020901 for dataset in which modification mortgages are deleted (-0.0206478 and -0.020901 are 25th percentiles of Call option value in two dataset respectively)	Call option value below -0.0213842 for dataset which is right censored and Call option value below -0.0214279 for dataset in which modification mortgages are deleted (-0.0213842 and -0.0214279 are 25th percentiles of Call option value in two dataset respectively)
Call option part 2	Call option value is from -0.0206478 to 0.0436119 for dataset which is right censored and Call option value is from -0.020901 to 0.0438131 for dataset in which modification mortgages are deleted (0.0436119 and 0.0438131 are 50th percentiles of Call option value in two dataset respectively)	Call option value is from -0.0213842 to 0.0422426 for dataset which is right censored and Call option value is from -0.0214279 to 0.0425066 for dataset in which modification mortgages are deleted (0.0422426 and 0.0425066 are 50th percentiles of Call option value in two dataset respectively)
Call option part 3	Call option value is from 0.0436119 to 0.1110333 for dataset which is right censored and Call option value is from 0.0438131 to 0.1111361 for dataset in which modification mortgages are deleted (0.1110333 and 0.1111361 are 75th percentiles of Call option value in two dataset respectively)	Call option value is from 0.0422426 to 0.109703 for dataset which is right censored and Call option value is from 0.0425066 to 0.1098711 for dataset in which modification mortgages are deleted (0.109703 and 0.1098711 are 75th percentiles of Call option value in two dataset respectively)
Call option part 4	Call option value above 0.1110333 for dataset which is right censored and Call option value above 0.1111361 for dataset in which modification mortgages are deleted	Call option value above 0.109703 for dataset which is right censored and Call option value above 0.1098711 for dataset in which modification mortgages are deleted
The unemployment rate variables		
The unemployment rate part 1	The unemployment rate in Tampa below 5.8 (5.8 is 50th percentile of the unemployment rate in both datasets)	The unemployment rate in Tampa below 5.8 (5.8 is 50th percentile of the unemployment rate in both datasets)
The unemployment rate part 2	The unemployment rate in Tampa above 5.8	The unemployment rate in Tampa above 5.8
Negative equity variables		
Negative equity part 1	Negative equity value below 10.04659 for dataset which is right censored and Negative equity value below 10.05241 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 10.04659 and 10.05241 are 25th percentiles of	Negative equity value below 9.953914 for dataset which is right censored and Negative equity value below 9.96664 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 9.953914 and 9.96664 are 25th percentiles of

	negative equity value in two dataset respectively)	negative equity value in two dataset respectively)
Negative equity part 2	Negative equity value is from 10.04659 to 22.00146 for dataset which is right censored and Negative equity value is from 10.05241 to 22.0039 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 22.00146 and 22.0039 are 50^{th} percentiles of negative equity value in two dataset respectively)	Negative equity value is from 9.953914 to 21.95684 for dataset which is right censored and Negative equity value is from 9.96664 to 21.95579 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 21.95684 and 21.95579 are 50^{th} percentiles of negative equity value in two dataset respectively)
Negative equity part 3	Negative equity value is from 22.00146 to 37.18636 for dataset which is right censored and Negative equity value is from 22.0039 to 37.11626 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 37.18636 and 37.11626 are 75^{th} percentiles of negative equity value in two dataset respectively)	Negative equity value is from 21.95684 to 37.20905 for dataset which is right censored and Negative equity value is from 21.95579 to 37.14248 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 37.20905 and 37.14248 are 75^{th} percentiles of negative equity value in two dataset respectively)
Negative equity part 4	Negative equity value above 37.18636 for dataset which is right censored and Negative equity value above 37.11626 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)	Negative equity value above 37.20905 for dataset which is right censored and Negative equity value above 37.14248 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)
The negative equity dummy	Equal to 1 if the equity less than zero in that month and 0 otherwise.	Equal to 1 if the equity less than zero in that month and 0 otherwise.
Original loan-to-value variables		
Original LTV part 1	Original loan-to-value below 0.67 (Original loan-to-value is divided by 100; 0.67 is 25^{th} percentiles of original loan-to-value in both datasets)	Original loan-to-value below 0.67 for dataset which is right censored and below 0.66 for dataset in which modification mortgages are deleted (Original loan-to-value is divided by 100; 0.67 and 0.66 are 25^{th} percentiles of original loan-to-value in two dataset respectively)
Original LTV part 2	Original loan-to-value is from 0.67 to 0.8 (0.8 is 50^{th} percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.67 to 0.8 for dataset which is right censored and from 0.66 to 0.8 for dataset in which modification mortgages are deleted (0.8 is 50^{th} percentiles of original loan-to-value in both datasets)
Original LTV part 3	Original loan-to-value is from 0.8 to 0.95 (0.80 is 75^{th} percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.8 to 0.95 (0.95 is 75^{th} percentiles of original loan-to-value in both datasets)
Original LTV part 4	Original loan-to-value above 0.95	Original loan-to-value above 0.95
Log loan size variables		
Log loan size part 1	Log loan size value is taking log of loan size. Log loan size value below 11.40757 (11.40757 is 25^{th} percentiles of log loan size value in both datasets)	Log loan size value is taking log of loan size. Log loan size value below 11.40757 (11.40757 is 25^{th} percentiles of log loan size value in both datasets)
Log loan size part 2	Log loan size value is from 11.40757 to 11.77529 (11.77529 is 50^{th} percentiles of log loan size value in both datasets)	Log loan size value is from 11.40757 to 11.77529 (11.77529 is 50^{th} percentiles of log loan size value in both datasets)
Log loan size part 3	Log loan size value is from 11.77529 to 12.11176 (12.11176 is 75^{th} percentiles of log loan size value in both datasets)	Log loan size value is from 11.77529 to 12.11176 (12.11176 is 75^{th} percentiles of log loan size value in both datasets)
Log loan size part 4	Log loan size value above 12.11176	Log loan size value above 12.11176
Loan age variables		

Loan age part 1	Loan age value below 13 (13 is 25th percentiles of loan age value in both datasets)	Loan age value below 12 (12 is 25th percentiles of loan age value in both datasets)
Loan age part 2	Loan age value is from 13 to 28 (28 is 50th percentiles of loan age value in both datasets)	Loan age value is from 12 to 27 (27 is 50th percentiles of loan age value in both datasets)
Loan age part 3	Loan age value is from 28 to 51 (51 is 75th percentiles of loan age value in both datasets)	Loan age value is from 27 to 51 (51 is 75th percentiles of loan age value in both datasets)
Loan age part 4	Loan age value above 51	Loan age value above 51
The credit score variables		
The credit score part 1	The credit score value below 0.677 (The credit score value is divided by 1000; 0.677 is 25th percentiles of the credit score value in both datasets)	The credit score value below 0.677 for dataset which is right censored and The credit score value below 0.678 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.677 and 0.678 are 25th percentiles of the credit score value in two dataset respectively)
The credit score part 2	The credit score value is from 0.677 to 0.723 for dataset which is right censored and 0.724 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.723 and 0.724 are 50th percentiles of the credit score value in two dataset respectively)	The credit score value is from 0.677 to 0.724 for dataset which is right censored and from 0.678 to 0.725 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.724 and 0.725 are 50th percentiles of the credit score value in two dataset respectively)
The credit score part 3	The credit score value is from 0.723 to 0.767 for dataset which is right censored and from 0.724 to 0.767 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.767 is 75th percentiles of the credit score value in both datasets)	The credit score value is from 0.724 to 0.767 for dataset which is right censored and from 0.725 to 0.768 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.767 and 0.768 are 75th percentiles of the credit score value in two dataset respectively)
The credit score part 4	The credit score value above 0.767 (The credit score value is divided by 1000)	The credit score value above 0.767 for dataset which is right censored and above 0.768 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000)
The debt-to-income ratio variables		
DTI part 1	The debt-to-income ratio value below 0.26 (The debt-to-income ratio value is divided by 100; 0.26 is 25th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value below 0.26 (The debt-to-income ratio value is divided by 100; 0.26 is 25th percentiles of the debt-to-income ratio value in both datasets)
DTI part 2	The debt-to-income ratio value is from 0.26 to 0.35 (The debt-to-income ratio value is divided by 100; 0.35 is 50th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.26 to 0.35 (The debt-to-income ratio value is divided by 100; 0.35 is 50th percentiles of the debt-to-income ratio value in both datasets)
DTI part 3	The debt-to-income ratio value is from 0.35 to 0.43 (The debt-to-income ratio value is divided by 100; 0.43 is 75th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.35 to 0.43 (The debt-to-income ratio value is divided by 100; 0.43 is 75th percentiles of the debt-to-income ratio value in both datasets)
DTI part 4	The debt-to-income ratio value above 0.43 (The debt-to-income ratio value is divided by 100)	The debt-to-income ratio value above 0.43 (The debt-to-income ratio value is divided by 100)
The dummy for the number of units	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1
Time period variables		

Time period part 1	Time period starts from March 1999 to July 2003 for prepayment and time period start from March 1999 to February 2008 for default	Time period starts from March 1999 to July 2003 for prepayment and time period start from March 1999 to February 2008 for default
Time period part 2	Time period starts from July 2003 to November 2008 for prepayment and time period start from February 2008 to May 2009 for default	Time period starts from July 2003 to November 2008 for prepayment and time period start from February 2008 to May 2009 for default
Time period part 3	Time period starts from November 2008 to March 2013 for prepayment and time period start from May 2009 to March 2013 for default	Time period starts from November 2008 to March 2013 for prepayment and time period start from May 2009 to March 2013 for default

Table 13. Explanatory Variables Definition for Detroit

Variable Name	Variable Definition for Different Dependent Variables	
	Prepayment and default	Prepayment and 90-days-delinquency
Prepayment penalty	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.
Call option variables		
Call option part 1	Call option value below -0.0236173 for dataset which is right censored and Call option value below -0.023294 for dataset in which modification mortgages are deleted (-0.0236173 and -0.023294 are 25th percentiles of Call option value in two dataset respectively)	Call option value below -0.0249068 for dataset which is right censored and Call option value below -0.0246977 for dataset in which modification mortgages are deleted (-0.0249068 and -0.0246977 are 25th percentiles of Call option value in two dataset respectively)
Call option part 2	Call option value is from -0.0236173 to 0.0383375 for dataset which is right censored and Call option value is from -0.023294 to 0.0385326 for dataset in which modification mortgages are deleted (0.0383375 and 0.0385326 are 50th percentiles of Call option value in two dataset respectively)	Call option value is from -0.0249068 to 0.0365674 for dataset which is right censored and Call option value is from -0.0246977 to 0.0368146 for dataset in which modification mortgages are deleted (0.0365674 and 0.0368146 are 50th percentiles of Call option value in two dataset respectively)
Call option part 3	Call option value is from 0.0383375 to 0.1080026 for dataset which is right censored and Call option value is from 0.0385326 to 0.108192 for dataset in which modification mortgages are deleted (0.1080026 and 0.108192 are 75th percentiles of Call option value in two dataset respectively)	Call option value is from 0.0365674 to 0.1057105 for dataset which is right censored and Call option value is from 0.0368146 to 0.1060208 for dataset in which modification mortgages are deleted (0.1057105 and 0.1060208 are 75th percentiles of Call option value in two dataset respectively)
Call option part 4	Call option value above 0.1080026 for dataset which is right censored and Call option value above 0.108192 for dataset in which modification mortgages are deleted	Call option value above 0.1057105 for dataset which is right censored and Call option value above 0.1060208 for dataset in which modification mortgages are deleted
The unemployment rate variables		
The unemployment rate part 1	The unemployment rate in Detroit below 7.900001 (7.900001 is 50th percentile of the unemployment rate in both datasets)	The unemployment rate in Detroit below 7.900001 (7.900001 is 50th percentile of the unemployment rate in both datasets)
The unemployment rate part 2	The unemployment rate in Detroit above 7.900001	The unemployment rate in Detroit above 7.900001
Negative equity variables		
Negative equity part 1	Negative equity value below 8.397582 for dataset which is right censored and Negative	Negative equity value below 8.39557 for dataset which is right censored and Negative

	equity value below 8.382367 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 8.397582 and 8.382367 are 25th percentiles of negative equity value in two dataset respectively)	equity value below 8.385215 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 8.39557 and 8.385215 are 25th percentiles of negative equity value in two dataset respectively)
Negative equity part 2	Negative equity value is from 8.397582 to 17.91852 for dataset which is right censored and Negative equity value is from 8.382367 to 17.83435 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 17.91852 and 17.83435 are 50th percentiles of negative equity value in two dataset respectively)	Negative equity value is from 8.39557 to 17.92759 for dataset which is right censored and Negative equity value is from 8.385215 to 17.8463 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 17.92759 and 17.8463 are 50th percentiles of negative equity value in two dataset respectively)
Negative equity part 3	Negative equity value is from 17.91852 to 30.10155 for dataset which is right censored and Negative equity value is from 17.83435 to 29.93952 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 30.10155 and 29.93952 are 75th percentiles of negative equity value in two dataset respectively)	Negative equity value is from 17.92759 to 30.06783 for dataset which is right censored and Negative equity value is from 17.8463 to 29.88266 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 30.06783 and 29.88266 are 75th percentiles of negative equity value in two dataset respectively)
Negative equity part 4	Negative equity value above 30.10155 for dataset which is right censored and Negative equity value above 29.93952 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)	Negative equity value above 30.06783 for dataset which is right censored and Negative equity value above 29.88266 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)
The negative equity dummy	Equal to 1 if the equity less than zero in that month and 0 otherwise.	Equal to 1 if the equity less than zero in that month and 0 otherwise.
Original loan-to-value variables		
Original LTV part 1	Original loan-to-value below 0.66 (Original loan-to-value is divided by 100; 0.66 is 25th percentiles of original loan-to-value in both datasets)	Original loan-to-value below 0.65 (Original loan-to-value is divided by 100; 0.65 is 25th percentiles of original loan-to-value in both datasets)
Original LTV part 2	Original loan-to-value is from 0.66 to 0.77 (Original loan-to-value is divided by 100; 0.77 is 50th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.65 to 0.77 (Original loan-to-value is divided by 100; 0.77 is 50th percentiles of original loan-to-value in both datasets)
Original LTV part 3	Original loan-to-value is from 0.77 to 0.80 (Original loan-to-value is divided by 100; 0.80 is 75th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.77 to 0.80 (Original loan-to-value is divided by 100; 0.80 is 75th percentiles of original loan-to-value in both datasets)
Original LTV part 4	Original loan-to-value is from 0.80 to 0.95 (Original loan-to-value is divided by 100; 0.95 is 95th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.80 to 0.95 (Original loan-to-value is divided by 100; 0.95 is 95th percentiles of original loan-to-value in both datasets)
Original LTV part 5	Original loan-to-value above 0.95	Original loan-to-value above 0.95
Log loan size variables		
Log loan size part 1	Log loan size value is taking log of loan size. Log loan size value below 11.28978 (11.28978 is 25th percentiles of log loan size value in both datasets)	Log loan size value is taking log of loan size. Log loan size value below 11.28978 (11.28978 is 25th percentiles of log loan size value in both datasets)
Log loan size part 2	Log loan size value is from 11.28978 to 11.63514 (11.63514 is 50th percentiles of log loan size value in both datasets)	Log loan size value is from 11.28978 to 11.63514 (11.63514 is 50th percentiles of log loan size value in both datasets)
Log loan size part 3	Log loan size value is from 11.63514 to 11.98293	Log loan size value is from 11.63514 to 11.98293

	(11.98293 is 75th percentiles of log loan size value in both datasets)	(11.98293 is 75th percentiles of log loan size value in both datasets)
Log loan size part 4	Log loan size value above 11.98293	Log loan size value above 11.98293
Loan age variables		
Loan age part 1	Loan age value below 13 (13 is 25th percentiles of loan age value in both datasets)	Loan age value below 13 (13 is 25th percentiles of loan age value in both datasets)
Loan age part 2	Loan age value is from 13 to 31 (31 is 50th percentiles of loan age value in both datasets)	Loan age value is from 13 to 30 (30 is 50th percentiles of loan age value in both datasets)
Loan age part 3	Loan age value is from 31 to 58 (58 is 75th percentiles of loan age value in both datasets)	Loan age value is from 30 to 57 (57 is 75th percentiles of loan age value in both datasets)
Loan age part 4	Loan age value above 58	Loan age value above 57
The credit score variables		
The credit score part 1	The credit score value below 0.674 for dataset which is right censored and The credit score value below 0.675 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.674 and 0.675 are 25th percentiles of the credit score value in two dataset respectively)	The credit score value below 0.676 (The credit score value is divided by 1000; 0.676 is 25th percentiles of the credit score value in both datasets)
The credit score part 2	The credit score value is from 0.674 to 0.726 for dataset which is right censored and from 0.675 to 0.726 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.726 is 50th percentiles of the credit score value in both datasets)	The credit score value is from 0.676 to 0.727 for dataset which is right censored and to 0.728 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.727 and 0.728 are 50th percentiles of the credit score value in two dataset respectively)
The credit score part 3	The credit score value is from 0.726 to 0.768 for dataset which is right censored and to 0.769 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.768 and 0.769 are 75th percentiles of the credit score value in two dataset respectively)	The credit score value is from 0.727 to 0.769 for dataset which is right censored and from 0.728 to 0.769 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.769 is 75th percentiles of the credit score value in both datasets)
The credit score part 4	The credit score value above 0.768 for dataset which is right censored and above 0.769 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000)	The credit score value above 0.769 (The credit score value is divided by 1000)
The debt-to-income ratio variables		
DTI part 1	The debt-to-income ratio value below 0.24 (The debt-to-income ratio value is divided by 100; 0.24 is 25th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value below 0.24 (The debt-to-income ratio value is divided by 100; 0.24 is 25th percentiles of the debt-to-income ratio value in both datasets)
DTI part 2	The debt-to-income ratio value is from 0.24 to 0.32 (The debt-to-income ratio value is divided by 100; 0.32 is 50th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.24 to 0.32 (The debt-to-income ratio value is divided by 100; 0.32 is 50th percentiles of the debt-to-income ratio value in both datasets)
DTI part 3	The debt-to-income ratio value is from 0.32 to 0.41 (The debt-to-income ratio value is divided by 100; 0.41 is 75th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.32 to 0.41 (The debt-to-income ratio value is divided by 100; 0.41 is 75th percentiles of the debt-to-income ratio value in both datasets)
DTI part 4	The debt-to-income ratio value above 0.41 (The debt-to-income ratio value is divided by 100)	The debt-to-income ratio value above 0.41 (The debt-to-income ratio value is divided by 100)

The dummy for the number of units	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1
Time period variables		
Time period part 1	Time period starts from March 1999 to July 2003 for prepayment and time period start from March 1999 to July 2008 for default	Time period starts from March 1999 to July 2003 for prepayment and time period start from March 1999 to July 2008 for default
Time period part 2	Time period starts from July 2003 to November 2008 for prepayment and time period start from to July 2008 to January 2010 for default	Time period starts from July 2003 to November 2008 for prepayment and time period start from to July 2008 to January 2010 for default
Time period part 3	Time period starts from November 2008 to March 2013 for prepayment and time period start from January 2010 to March 2013 for default	Time period starts from November 2008 to March 2013 for prepayment and time period start from January 2010 to March 2013 for default

Table 14. Explanatory Variables Definition for Las Vegas

Variable Name	Variable Definition for Different Dependent Variables	
	Prepayment and default	Prepayment and 90-days-delinquency
Prepayment penalty	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.	A dummy variable, which equal to 1 when borrowers have to pay additional certain amount of principal to get prepayment, and equal to 0 otherwise.
Call option variables		
Call option part 1	Call option value below -0.0226132 for dataset which is right censored and Call option value below -0.0221891 for dataset in which modification mortgages are deleted (-0.0226132 and -0.0221891 are 25th percentiles of Call option value in two dataset respectively)	Call option value below -0.0237949 for dataset which is right censored and Call option value below -0.0221891 for dataset in which modification mortgages are deleted (-0.0237949 and -0.0221891 are 25th percentiles of Call option value in two dataset respectively)
Call option part 2	Call option value is from -0.0226132 to 0.0408946 for dataset which is right censored and Call option value is from -0.0221891 to 0.0414338 for dataset in which modification mortgages are deleted (0.0408946 and 0.0414338 are 50th percentiles of Call option value in two dataset respectively)	Call option value is from -0.0237949 to 0.0393865 for dataset which is right censored and Call option value is from -0.0221891 to 0.0414338 for dataset in which modification mortgages are deleted (0.0393865 and 0.0414338 are 50th percentiles of Call option value in two dataset respectively)
Call option part 3	Call option value is from 0.0408946 to 0.1056783 for dataset which is right censored and Call option value is from 0.0414338 to 0.1060248 for dataset in which modification mortgages are deleted (0.1056783 and 0.1060248 are 75th percentiles of Call option value in two dataset respectively)	Call option value is from 0.0393865 to 0.1041278 for dataset which is right censored and Call option value is from 0.0414338 to 0.1060248 for dataset in which modification mortgages are deleted (0.1041278 and 0.1060248 are 75th percentiles of Call option value in two dataset respectively)
Call option part 4	Call option value above 0.1056783 for dataset which is right censored and Call option value above 0.1060248 for dataset in which modification mortgages are deleted	Call option value above 0.1041278 for dataset which is right censored and Call option value above 0.1060248 for dataset in which modification mortgages are deleted
The unemployment rate variables		
The unemployment rate part 1	The unemployment rate in Las Vegas below 6.300001 (6.300001 is 50th percentile of the unemployment rate in both datasets)	The unemployment rate in Las Vegas below 6.2 (6.2 is 50th percentile of the unemployment rate in both datasets)
The unemployment rate part 2	The unemployment rate in Las Vegas above 6.300001	The unemployment rate in Las Vegas above 6.2

Negative equity variables		
Negative equity part 1	Negative equity value below 21.45515 for dataset which is right censored and Negative equity value below 21.19559 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 21.45515 and 21.19559 are 25th percentiles of negative equity value in two dataset respectively)	Negative equity value below 20.99858 for dataset which is right censored and Negative equity value below 20.72638 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 20.99858 and 20.72638 are 25th percentiles of negative equity value in two dataset respectively)
Negative equity part 2	Negative equity value is from 21.45515 to 44.30557 for dataset which is right censored and Negative equity value is from 21.19559 to 43.88683 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 44.30557 and 43.88683 are 50th percentiles of negative equity value in two dataset respectively)	Negative equity value is from 20.99858 to 43.57627 for dataset which is right censored and Negative equity value is from 20.72638 to 43.18412 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 43.57627 and 43.18412 are 50th percentiles of negative equity value in two dataset respectively)
Negative equity part 3	Negative equity value is from 44.30557 to 74.92264 for dataset which is right censored and Negative equity value is from 43.88683 to 74.37087 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 74.92264 and 74.37087 are 75th percentiles of negative equity value in two dataset respectively)	Negative equity value is from 43.57627 to 74.12209 for dataset which is right censored and Negative equity value is from 43.18412 to 73.44044 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000; 74.12209 and 73.44044 are 75th percentiles of negative equity value in two dataset respectively)
Negative equity part 4	Negative equity value above 74.92264 for dataset which is right censored and Negative equity value above 74.37087 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)	Negative equity value above 74.12209 for dataset which is right censored and Negative equity value above 73.44044 for dataset in which modification mortgages are deleted (Negative equity value is divided by 1000)
The negative equity dummy	Equal to 1 if the equity less than zero in that month and 0 otherwise.	Equal to 1 if the equity less than zero in that month and 0 otherwise.
Original loan-to-value variables		
Original LTV part 1	Original loan-to-value below 0.65 (Original loan-to-value is divided by 100; 0.65 is 25th percentiles of original loan-to-value in both datasets)	Original loan-to-value below 0.65 (Original loan-to-value is divided by 100; 0.65 is 25th percentiles of original loan-to-value in both datasets)
Original LTV part 2	Original loan-to-value is from 0.65 to 0.78 (Original loan-to-value is divided by 100; 0.78 is 50th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.65 to 0.78 (Original loan-to-value is divided by 100; 0.78 is 50th percentiles of original loan-to-value in both datasets)
Original LTV part 3	Original loan-to-value is from 0.78 to 0.80 (Original loan-to-value is divided by 100; 0.80 is 75th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.78 to 0.80 (Original loan-to-value is divided by 100; 0.80 is 75th percentiles of original loan-to-value in both datasets)
Original LTV part 4	Original loan-to-value is from 0.80 to 0.95 (Original loan-to-value is divided by 100; 0.95 is 95th percentiles of original loan-to-value in both datasets)	Original loan-to-value is from 0.80 to 0.95 (Original loan-to-value is divided by 100; 0.95 is 95th percentiles of original loan-to-value in both datasets)
Original LTV part 5	Original loan-to-value above 0.95	Original loan-to-value above 0.95
Log loan size variables		
Log loan size part 1	Log loan size value is taking log of loan size. Log loan size value below 11.63514 for dataset which is right censored and Log loan size value below 11.62625 for dataset in which modification mortgages are deleted (11.63514 and 11.62625 are 25th percentiles of	Log loan size value is taking log of loan size. Log loan size value below 11.62625 (11.62625 is 25th percentiles of log loan size value in both datasets)

	log loan size value in both datasets)	
Log loan size part 2	Log loan size value is from 11.63514 to 11.97035 for dataset which is right censored and from 11.62625 to 11.964 for dataset in which modification mortgages are deleted (11.97035 and 11.964 are 50th percentiles of log loan size value in two dataset respectively)	Log loan size value is from 11.62625 to 11.97035 for dataset which is right censored and to 11.964 for dataset in which modification mortgages are deleted (11.97035 and 11.964 are 50th percentiles of log loan size value in two dataset respectively)
Log loan size part 3	Log loan size value is from 11.97035 to 12.30138 for dataset which is right censored and from 11.964 to 12.28765 for dataset in which modification mortgages are deleted (12.30138 and 12.28765 are 75th percentiles of log loan size value in two dataset respectively)	Log loan size value is from 11.97035 to 12.29683 for dataset which is right censored and from 11.964 to 12.28765 for dataset in which modification mortgages are deleted (12.29683 and 12.28765 are 75th percentiles of log loan size value in two dataset respectively)
Log loan size part 4	Log loan size value above 12.30138 for dataset which is right censored and Log loan size value above 12.28765 for dataset in which modification mortgages are deleted	Log loan size value above 12.29683 for dataset which is right censored and Log loan size value above 12.28765 for dataset in which modification mortgages are deleted
Loan age variables		
Loan age part 1	Loan age value below 11 (11 is 25th percentiles of loan age value in both datasets)	Loan age value below 11 (11 is 25th percentiles of loan age value in both datasets)
Loan age part 2	Loan age value is from 11 to 25 (25 is 50th percentiles of loan age value in both datasets)	Loan age value is from 11 to 25 (25 is 50th percentiles of loan age value in both datasets)
Loan age part 3	Loan age value is from 25 to 47 (47 is 75th percentiles of loan age value in both datasets)	Loan age value is from 25 to 46 for dataset which is right censored and to 47 for dataset in which modification mortgages are deleted (46 and 47 are 75th percentiles of loan age value in two dataset respectively)
Loan age part 4	Loan age value above 47	Loan age value above 46 for dataset which is right censored and above 47 for dataset in which modification mortgages are deleted
The credit score variables		
The credit score part 1	The credit score value below 0.684 for dataset which is right censored and The credit score value below 0.685 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.684 and 0.685 are 25th percentiles of the credit score value in two dataset respectively)	The credit score value below 0.684 for dataset which is right censored and The credit score value below 0.685 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.684 and 0.685 are 25th percentiles of the credit score value in two dataset respectively)
The credit score part 2	The credit score value is from 0.684 to 0.731 for dataset which is right censored and from 0.685 to 0.731 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.731 is 50th percentiles of the credit score value in both datasets)	The credit score value is from 0.684 to 0.731 for dataset which is right censored and from 0.685 to 0.731 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.731 is 50th percentiles of the credit score value in both datasets)
The credit score part 3	The credit score value is from 0.731 to 0.771 for dataset which is right censored and to 0.772 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000; 0.771 and 0.772 are 75th percentiles of the credit score value in two dataset respectively)	The credit score value is from 0.731 to 0.772 (The credit score value is divided by 1000; 0.772 is 75th percentiles of the credit score value in both datasets)
The credit score part 4	The credit score value above 0.771 for dataset which is right censored and above 0.772 for dataset in which modification mortgages are deleted (The credit score value is divided by 1000)	The credit score value above 0.772 (The credit score value is divided by 1000)
The debt-to-income ratio		

variables		
DTI part 1	The debt-to-income ratio value below 0.27 (The debt-to-income ratio value is divided by 100; 0.27 is 25th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value below 0.27 (The debt-to-income ratio value is divided by 100; 0.27 is 25th percentiles of the debt-to-income ratio value in both datasets)
DTI part 2	The debt-to-income ratio value is from 0.27 to 0.36 (The debt-to-income ratio value is divided by 100; 0.36 is 50th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.27 to 0.36 (The debt-to-income ratio value is divided by 100; 0.36 is 50th percentiles of the debt-to-income ratio value in both datasets)
DTI part 3	The debt-to-income ratio value is from 0.36 to 0.45 (The debt-to-income ratio value is divided by 100; 0.45 is 75th percentiles of the debt-to-income ratio value in both datasets)	The debt-to-income ratio value is from 0.36 to 0.45 (The debt-to-income ratio value is divided by 100; 0.45 is 75th percentiles of the debt-to-income ratio value in both datasets)
DTI part 4	The debt-to-income ratio value above 0.45 (The debt-to-income ratio value is divided by 100)	The debt-to-income ratio value above 0.45 (The debt-to-income ratio value is divided by 100)
The dummy for the number of units	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1	This is a dummy variable which equal to 1 if property unit larger than 1 and equal to 0 if property unit equal to 1
Time period variables		
Time period part 1	Time period starts from March 1999 to August 2003 for prepayment and time period start from March 1999 to April 2008 for default	Time period starts from March 1999 to August 2003 for prepayment and time period start from March 1999 to April 2008 for default
Time period part 2	Time period starts from August 2003 to November 2008 for prepayment and time period start from April 2008 to June 2010 for default	Time period starts from August 2003 to November 2008 for prepayment and time period start from April 2008 to June 2010 for default
Time period part 3	Time period starts from November 2008 to March 2013 for prepayment and time period start from June 2010 to December 2011 for default	Time period starts from November 2008 to March 2013 for prepayment and time period start from June 2010 to December 2011 for default
Time period part 4	Time period start from December 2011 to March 2013 for default	Time period start from December 2011 to March 2013 for default

Table 9 shows that, for Phoenix, the mean value of the call option is about 0.039; for Miami, the mean value of the call option is about 0.049; for Tampa, the mean value of the call option is about 0.044; for Detroit, the mean value of the call option is about 0.042; and for Las Vegas, the mean value of the call option is about 0.041. To capture the nonlinear relationship between the value of the call option and prepayment, a linear spline is used. The knots of the linear spline are from the first, second and third quartile of the value of the call option in each MSA. For example, with Phoenix, the knots are -0.025, 0.037 and 0.102 for both datasets. For Miami, the knots are -0.017, 0.048 and 0.115 for the dataset with right censored mortgages and are -0.016, 0.049 and 0.115 for the dataset with modified mortgages deleted.

The monthly unemployment rates are merged into the individual MSA dataset to explain the default risk. The unemployment status of the households directly affects their ability to keep the mortgages. Therefore, hypothesis 3 argues:

H3: More defaults occur in months with a higher unemployment rate.

For Phoenix, the mean value of the unemployment rate is about 6.202; for Miami, the mean value is about 6.916; for Tampa, the mean value is about 7.085; for Detroit, the mean value is about 9.105; for Las Vegas, the mean value is about 8.417, respectively. This indicates the unemployment rate in Detroit is much higher than that in other MSAs. A linear spline is used to capture the nonlinear relationship and the knot is the second quartile of the unemployment rate. For example, for Phoenix, the knot is 5.7 for both datasets. For Miami, the knot is 6.2 for the dataset with right censored mortgage and is 6.1 for the dataset with modified mortgages deleted.

Three financial explanatory variables are used to analyze the default risk. They are negative equity, a dummy for negative equity and original loan-to-value. Their definitions are presented first followed by a discussion regarding how these explanatory variables are used.

Negative equity is the difference between the market value of a property and the remaining balance of the mortgage. The calculation process is presented in the Appendix A. When the current housing value is less than the remaining balance (or the value) of the mortgage, so called the put option is "in the money," households have more incentive to default. Therefore, hypothesis 4 suggests:

H4: There is a positive relationship between negative equity and default risk.

The mean value of negative equity is calculated conditional on the negative equity dummy equal to one. For Phoenix, the mean value is about 8.704; for Miami, the mean value is about 5.908; for Tampa, the mean value is about 4.570; for Detroit, the mean value is about 5.244; for Las Vegas, the mean value is about 13.294. Among all MSAs, Las Vegas has the highest negative equity. A linear spline is used for negative equity and the knots are from the first, second and third quartile of the non-zero negative equity value. For example, with Phoenix, the knots are 18.366, 37.804 and 62.910 for the dataset with right censored mortgages, and are 18.256, 37.613 and 62.599 for the dataset with modified mortgages deleted. For Miami, the knots are 13.194, 30.568 and 53.854 for the dataset with right censored mortgages, and are 13.189, 30.443 and 53.756 for the dataset with modified mortgages deleted.

The dummy variable for negative equity equals to one if the equity in a month is negative. Otherwise the variable equals to zero. The conjecture will be:

H5: There is a positive coefficient for the negative equity dummy.

The effect of original loan-to-value on the default risk is shown to be positive in previous studies (eg. Deng et al 2000); therefore, the conjecture is:

H6: The original loan-to-value positively affects the default risk.

For Phoenix, the mean value of original loan-to-value is about 0.735; for Miami; the mean value is about 0.741; for Tampa, the mean value is about 0.744; for Detroit, the mean value is about 0.741; for Las Vegas, the mean value is about 0.749. Original loan-to-value in this paper's model is separated by four knots, which are the 25^{th}, the 50^{th}, the 75^{th}, and the 95^{th} percentile of its value. For example, with Phoenix, the knots are 0.65, 0.78, 0.80 and 0.95 for both datasets. For Miami, the knots are 0.63, 0.77, 0.80 and 0.95 for both datasets.

These three explanatory variables are combined into seven groups[7] to explain the default risk; the best three combinations in each MSA are listed in the results part.

This paper also enquires about the effect of the credit score and the debt-to-income ratio on the competing risks of prepayment and default. Theoretically, the households with higher credit score are less likely to default. Therefore, the conjecture is:

H7: The effect of the credit score on the default risk is negative.

However, whether or not the credit score affects prepayment risk is ambiguous. Accordingly, two estimations to explain the effect of the credit score on the prepayment and default risks are run. One estimation considers how the credit score affects both termination risks and the other estimation considers how the credit score affects only default, and the effect on prepayment is constrained to be zero. A linear spline for the credit score is used in both models

[7] The seven groups are: negative equity only, the negative equity dummy only, original loan-to-value only, negative equity and the negative equity dummy, negative equity and original loan-to-value, the negative equity dummy and original loan-to-value, and negative equity, the negative equity dummy together with original loan-to-value.

and the knots are the first, second and third quartile of the credit score. For example, for Phoenix, the knots are 0.688, 0.736 and 0.774 for the dataset with right censored mortgages, and are 0.688, 0.737 and 0.774 for the dataset with modified mortgages deleted. For Miami, the knots are 0.671, 0.71 and 0.753 for the dataset with right censored mortgages, and are 0.672, 0.711 and 0.754 for the dataset with modified mortgages deleted. Table 9 shows the mean value of the credit score in each MSA. For Phoenix, the mean value of the credit score is about 0.730; for Miami, the mean value of the credit score is about 0.708; for Tampa, the mean value is about 0.721; for Detroit, the mean value is about 0.718; for Las Vegas, it is about 0.726.

The debt-to-income ratio is an important measure of the ability to manage the payments. A large debt-to-income ratio indicates households spend a large part of their income to pay back the debt and they use a small amount of money to purchase other utilities. Therefore, the default risk is higher for households who have a larger debt-to-income ratio. The conjecture is:

H8: The relationship between the debt-to-income ratio and the default risk is positive.

However, whether the smaller debt-to-income ratio will lead to a larger possibility of prepayment is ambiguous. A linear spline is used to analyze the nonlinear relationship between the debt-to-income ratio and the termination risks. The knots form the first, second and third quartile of the ratio. For example, for Phoenix, the knots are 0.25, 0.34 and 0.42 for both datasets. For Miami, the knots are 0.25, 0.35 and 0.43 for the dataset with right censored mortgages and are 0.25, 0.34 and 0.43 for the dataset with modified mortgages deleted. Table 9 shows the mean value of the debt-to-income ratio. For Phoenix, the mean value of the debt-to-income ratio is about 0.341; for Miami, the mean value is about 0.349; for Tampa, the mean value is about 0.346; for Detroit, the mean value is about 0.334; for Las Vegas, it is about 0.357.

In this paper, the value for log loan size, loan age month[8] and The dummy for the number of units (the dummy variable equals to one if house units are more than one and equals to zero otherwise) are controlled. A linear spline for log loan size and loan age month is also used in both models and the knots come from the first, second and third quartile. The mean value, standard deviation, minimum and maximum values of these expletory variables are listed in table 9.

In this study, the explanatory variables are separated into six different groups. These six groups come from the effect of the credit score on both termination risks and only on the default risk, together with the best three combinations of negative equity, the negative equity dummy and original loan-to-value. To make the analysis more comprehensive, two different combinations of dependent variables are analyzed. The first combination contains the prepayment and the default and the second contains the prepayment and the 90-days-delinquency[9]. Here, 90-days-delinquency is a commonly used benchmark in the mortgage industry as an early warning indicator of default.

The econometric model and results for each MSA are introduced in the following section.

Econometric Methodology

The single-family mortgage data is monthly, therefore a discrete-time model is used to analyze the competing risks of prepayment and default.

Let T_i denote the termination month for mortgage i. No termination month can exceed the number of months in the term of a mortgage or a censoring month, in cases where the

[8] Here, it is assumed mortgages start from the first day of first payment month and the first payment month is treated as loan age 1, the month after the first payment month is loan age 2, and so on. Therefore, a variable named Loan Age is created. This variable calculates the number of months mortgages survive on the market.
[9] Notice that not all loans in the status of 90-days-delinquency eventually default.

observation is no longer recorded but no default or prepayment has occurred (the observation can be censored, for example, if the mortgages has been renegotiated without prepayment). Let k_i denote the minimum of the number of months in the term of the mortgage and the censoring month. Therefore,

$$T_i = \min({T_i}^D, {T_i}^P, k_i), \qquad 1.1$$

where ${T_i}^D$ denotes the random termination month of default for mortgage i, and ${T_i}^p$denotes the random termination month of prepayment for mortgage i. Let $Z_i(t)$ be the d-dimensional covariate vector that influences the termination decision for the mortgagor at time t. Finally, let $\delta_{Di}(t)$ be the indicator for whether the mortgage defaults at month t, and let $\delta_{Pi}(t)$ be the indicator for whether the mortgage prepays at month t.

The conditional probability of a default at time t (survival past time $t-1$) for mortgage i is defined to be:

$$\begin{aligned} {p_i}^D(t) &= \Pr\left({T_i}^D < t | {T_i}^D \geq t-1, Z_i(t)\right), && t = 1 \ldots T_i \\ &= \frac{\exp(\alpha^D + Z_i(t)\beta^D)}{1 + \exp(\alpha^D + Z_i(t)\beta^D) + \exp(\alpha^P + Z_i(t)\beta^P)}, && t = 1 \ldots T_i \qquad 1.2 \end{aligned}$$

where α^D and β^D are default parameters. The conditional probability of a prepayment at time t (survival past time $t-1$) is given by:

$$\begin{aligned} {p_i}^P(t) &= \Pr({T_i}^P < t / {T_i}^P \geq t-1, Z_i(t)), && t = 1 \ldots T_i \\ &= \frac{\exp(\alpha^P + Z_i(t)\beta^P)}{1 + \exp(\alpha^D + Z_i(t)\beta^D) + \exp(\alpha^P + Z_i(t)\beta^P)}, && t = 1 \ldots T_i \qquad 1.3 \end{aligned}$$

where α^P and β^P are prepayment parameters. The conditional probability of continuing with the mortgage at time t (survival past time $t-1$) i is defined to be:

$$p_i{}^C(t) = \frac{1}{1+\exp(\alpha^D + Z_i(t)\beta^D) + \exp(\alpha^P + Z_i(t)\beta^P)}, \qquad t = 1 \ldots T_i \qquad 1.4$$

where α^D, β^D, α^P and β^P are the parameters vector to be estimated.

Note that a default event and a prepayment event cannot be both observed in the same time period for any observations, since once one of the events occurs the observation ends. The likelihood contribution for mortgage i is now given by:

$$L_i = \prod_{t=1}^{T_i} (p_i{}^D(t)^{\delta_{Di}(t)} \times p_i{}^P(t)^{\delta_{Pi}(t)} \times p_i{}^C(t)^{1-\delta_{Di}(t)-\delta_{Pi}(t)}), \qquad 1.5$$

Thus, the likelihood function for the sample is the product of likelihood contribution across all mortgages in the sample. This model is also used to analyze the competing risks of the prepayment and the 90-days-delinquency.

Each model is then evaluated using two different approaches. The first evaluation is how well models can predict the dependent variable. For this the Pseudo R-square is used as a measure of predictive power. The other evaluation is to assess the fitness of models; for this, the Bayesian information criterion (BIC) is used.

The Pseudo R-square will be calculated as:

$$R^2{}_{McF} = 1 - \frac{\ln(L_M)}{\ln(L_0)}, \qquad 1.6$$

and is also called McFadden's R-square. $\ln(L_\mathrm{M})$ is the estimated likelihood for the model with predictors and $\ln(L_0)$ is the estimated likelihood for the model without predictors.

The formula for the BIC used in this paper is:

$$BIC = LLH - \frac{1}{2} \times \ln(N) \times p \qquad 1.7$$

in which LLH is the log likelihood of a model, N is the number of mortgages used in the model and p is the number of parameters estimated, including the constant. Thus, the larger the BIC is, the better the model fits.

Results for Five MSAs and the Evaluation of the Models

In this section, the important results for each MSA are presented. This paper firstly explains the results for Phoenix in detail and then compares its results with those of other four MSAs to conclude that there is no national housing market.

Results for Phoenix

For Phoenix, twenty-four models are estimated. The estimations comprise two sets of dependent variables, the prepayment and the default being the first, and the prepayment and the 90-days-delinquency being the second. For both groups of dependent variables, there are two datasets. The first contains mortgages that are right censored one month before the modification and the second deletes all mortgages modified at any point in time. For each of these four cases, the best three specifications are chosen using BIC. For each of these models, two estimations are also run. The first estimation includes the effect of the credit score on both termination risks and the second includes the effect of the credit score only on the default/90-days-delinquency risk. Therefore, a total of twenty-four results are presented for Phoenix.

For Phoenix, when the prepayment and the default are used as dependent variables, the best specification includes negative equity, the negative equity dummy and original loan-to-value.

The second best specification includes negative equity and the negative equity dummy, and the third best specification includes the negative equity dummy and original loan-to-value. In the case where the prepayment and the 90-days-delinquency are used as the dependent variable, the best specification includes negative equity and the negative equity dummy. The second best includes negative equity and the third includes negative equity, the negative equity dummy and original loan-to-value.

Tables 15 through 24 list the results for Phoenix. The results are separated into two groups based on which combination of dependent variables is used.

Group 1: Prepayment and default are dependent variables

The first group of models includes negative equity, the negative equity dummy, and original loan-to-value as explanatory variables. The coefficients are listed in table 15 and the odds ratios are listed in table 16. The following explanations are based on the dataset with right censored mortgages. The results of the model with credit score for both termination risks are introduced first. They are then compared with the results of the model with the credit score only for the default risk. The results for the dataset with modified mortgages deleted can be found in tables 15 and 16.

[insert Table 15 and 16 are here]

In the model with the credit score used to explain both termination risks, the effect of the indicator for a prepayment penalty supports H1. The odds for mortgages with a prepayment penalty are only 0.444 times as high as the same odds for mortgages without a prepayment penalty.

Table 15. Coefficients for Phoenix: Prepayment and default are dependent variables, negative equity, original LTV and the negative equity dummy are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the default		The credit score only for default		The credit score for both the prepayment and the default		The credit score only for default	
	Prepayment	Default	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.813	--	-0.774	--	-0.816	--	-0.781	--
	(0.232)		(0.232)		(0.232)		(0.232)	
Call option variables								
Call option part 1	5.365	--	5.348	--	5.418	--	5.414	--
	(0.884)		(0.883)		(0.885)		(0.884)	
Call option part 2	7.492	--	7.337	--	7.558	--	7.406	--
	(0.892)		(0.892)		(0.891)		(0.890)	
Call option part 3	7.927	--	7.542	--	8.000	--	7.638	--
	(0.674)		(0.673)		(0.676)		(0.675)	
Call option part 4	1.892	--	1.314	--	2.001	--	1.451	--
	(0.427)		(0.425)		(0.428)		(0.425)	
The unemployment rate variables								
The unemployment rate part 1	--	0.271	--	0.272	--	0.279	--	0.279
		(0.104)		(0.104)		(0.104)		(0.104)
The unemployment rate part 2	--	0.111	--	0.111	--	0.112	--	0.112
		(0.058)		(0.058)		(0.058)		(0.058)
Negative equity variables								
Negative equity part 1	--	0.012	--	0.012	--	0.012	--	0.012
		(0.014)		(0.014)		(0.014)		(0.014)
Negative equity part 2	--	0.019	--	0.019	--	0.019	--	0.019
		(0.009)		(0.009)		(0.009)		(0.009)
Negative equity part 3	--	0.013	--	0.013	--	0.014	--	0.014
		(0.006)		(0.006)		(0.006)		(0.006)
Negative equity part 4	--	-0.000	--	-0.000	--	0.001	--	0.001
		(0.003)		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	0.934	--	0.934	--	0.938	--	0.937
		(0.211)		(0.211)		(0.212)		(0.212)
Original loan-to-value variables								
Original LTV part 1	--	3.230	--	3.230	--	3.327	--	3.327
		(1.172)		(1.172)		(1.176)		(1.176)
Original LTV part 2	--	1.738	--	1.738	--	1.664	--	1.665
		(1.227)		(1.227)		(1.228)		(1.228)
Original LTV part 3	--	-6.256	--	-6.256	--	-8.057	--	-8.058
		(5.353)		(5.354)		(5.355)		(5.356)
Original LTV part 4	--	3.369	--	3.370	--	3.578	--	3.578
		(0.754)		(0.754)		(0.753)		(0.753)
Original LTV part 5	--	2.420	--	2.416	--	2.359	--	2.352
		(4.562)		(4.562)		(4.585)		(4.585)
Log loan size variables								
Log loan size part 1	0.755	0.349	0.667	0.349	0.756	0.348	0.671	0.348
	(0.079)	(0.371)	(0.078)	(0.371)	(0.079)	(0.372)	(0.078)	(0.372)
Log loan size part 2	0.840	-0.459	0.809	-0.459	0.846	-0.421	0.817	-0.420
	(0.124)	(0.427)	(0.124)	(0.427)	(0.124)	(0.429)	(0.124)	(0.429)
Log loan size part 3	0.209	-0.134	0.198	-0.134	0.198	-0.133	0.188	-0.133
	(0.132)	(0.375)	(0.132)	(0.375)	(0.136)	(0.387)	(0.136)	(0.387)
Log loan size part 4	0.990	-0.030	1.001	-0.028	1.008	-0.052	1.018	-0.050
	(0.095)	(0.268)	(0.095)	(0.268)	(0.093)	(0.265)	(0.093)	(0.265)

Loan age variables								
Loan age part 1	0.140	0.419	0.140	0.419	0.140	0.418	0.140	0.418
	(0.008)	(0.116)	(0.008)	(0.116)	(0.008)	(0.116)	(0.008)	(0.116)
Loan age part 2	0.003	0.077	0.003	0.077	0.003	0.078	0.003	0.078
	(0.004)	(0.017)	(0.004)	(0.017)	(0.004)	(0.017)	(0.004)	(0.017)
Loan age part 3	-0.002	-0.006	-0.003	-0.006	-0.002	-0.007	-0.003	-0.007
	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)
Loan age part 4	0.006	-0.006	0.005	-0.006	0.006	-0.006	0.005	-0.006
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables								
The credit score part 1	1.801	-7.953	--	-7.980	1.764	-8.166	--	-8.193
	(0.682)	(1.481)		(1.481)	(0.679)	(1.475)		(1.475)
The credit score part 2	3.445	-5.924	--	-5.986	3.194	-6.380	--	-6.439
	(0.939)	(2.484)		(2.484)	(0.917)	(2.429)		(2.429)
The credit score part 3	0.993	-2.714	--	-2.732	0.921	-2.636	--	-2.653
	(1.232)	(3.907)		(3.907)	(1.264)	(4.015)		(4.015)
The credit score part 4	6.919	-20.685	--	-20.819	6.804	-20.754	--	-20.887
	(1.662)	(6.741)		(6.741)	(1.666)	(6.759)		(6.759)
The debt-to-income ratio variables								
DTI part 1	-0.870	1.154	-1.097	1.149	-0.863	1.184	-1.077	1.180
	(0.376)	(1.561)	(0.374)	(1.561)	(0.376)	(1.562)	(0.374)	(1.562)
DTI part 2	0.947	2.779	0.671	2.775	0.989	2.944	0.724	2.940
	(0.504)	(1.580)	(0.503)	(1.580)	(0.504)	(1.582)	(0.503)	(1.582)
DTI part 3	-0.665	0.744	-0.832	0.740	-0.561	0.898	-0.722	0.894
	(0.548)	(1.448)	(0.547)	(1.448)	(0.548)	(1.451)	(0.547)	(1.451)
DTI part 4	0.320	0.648	0.410	0.650	0.284	0.686	0.376	0.688
	(0.349)	(0.762)	(0.348)	(0.762)	(0.349)	(0.761)	(0.348)	(0.761)
The dummy for the number of units	-0.548	-0.531	-0.498	-0.530	-0.551	-0.558	-0.503	-0.558
	(0.227)	(0.712)	(0.227)	(0.712)	(0.227)	(0.712)	(0.227)	(0.712)
Time period variables								
Time period part 1	0.023	-0.004	0.023	-0.004	0.022	-0.003	0.023	-0.003
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	-0.037	0.036	-0.037	0.036	-0.037	0.034	-0.036	0.034
	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	0.017	0.004	0.019	0.004	0.016	0.004	0.018	0.004
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)
	Number of mortgages 12,734				Number of mortgages 12,556			

Table 16. Odds ratios for Phoenix: Prepayment and default are dependent variables, negative equity, original LTV and the negative equity dummy are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the default		The credit score only for default		The credit score for both the prepayment and the default		The credit score only for default	
	Prepayment	Default	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	0.444	--	0.461	--	0.442	--	0.458	--
	(0.103)		(0.107)		(0.102)		(0.106)	
Call option variables								
Call option part 1	213.758	--	210.201	--	225.394	--	224.591	--
	(189.015)		(185.706)		(199.393)		(198.510)	
Call option part 2	1,794.485	--	1,536.113	--	1,916.514	--	1,645.261	--

	(1,600.797)		(1,369.525)		(1,706.949)		(1,464.498)	
Call option part 3	2,770.925	--	1,885.200	--	2,980.178	--	2,074.706	--
	(1,868.153)		(1,268.629)		(2,014.044)		(1,399.589)	
Call option part 4	6.634	--	3.723	--	7.400	--	4.267	--
	(2.834)		(1.582)		(3.164)		(1.815)	
The unemployment rate variables								
The unemployment rate part 1	--	1.312	--	1.312	--	1.322	--	1.322
		(0.137)		(0.137)		(0.138)		(0.138)
The unemployment rate part 2	--	1.117	--	1.117	--	1.119	--	1.119
		(0.065)		(0.065)		(0.065)		(0.065)
Negative equity variables								
Negative equity part 1	--	1.012	--	1.012	--	1.012	--	1.012
		(0.014)		(0.014)		(0.014)		(0.014)
Negative equity part 2	--	1.019	--	1.019	--	1.020	--	1.020
		(0.009)		(0.009)		(0.009)		(0.009)
Negative equity part 3	--	1.014	--	1.014	--	1.014	--	1.014
		(0.006)		(0.006)		(0.006)		(0.006)
Negative equity part 4	--	1.000	--	1.000	--	1.001	--	1.001
		(0.003)		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	2.545	--	2.544	--	2.554	--	2.553
		(0.537)		(0.537)		(0.541)		(0.541)
Original loan-to-value variables								
Original LTV part 1	--	25.280	--	25.273	--	27.850	--	27.842
		(29.621)		(29.614)		(32.759)		(32.751)
Original LTV part 2	--	5.684	--	5.687	--	5.282	--	5.285
		(6.976)		(6.980)		(6.487)		(6.491)
Original LTV part 3	--	0.002	--	0.002	--	0.000	--	0.000
		(0.010)		(0.010)		(0.002)		(0.002)
Original LTV part 4	--	29.059	--	29.070	--	35.795	--	35.808
		(21.914)		(21.922)		(26.959)		(26.969)
Original LTV part 5	--	11.248	--	11.196	--	10.578	--	10.511
		(51.313)		(51.075)		(48.503)		(48.193)
Log loan size variables								
Log loan size part 1	2.128	1.418	1.949	1.418	2.130	1.416	1.956	1.416
	(0.168)	(0.526)	(0.152)	(0.526)	(0.168)	(0.526)	(0.153)	(0.526)
Log loan size part 2	2.317	0.632	2.246	0.632	2.330	0.657	2.264	0.657
	(0.287)	(0.270)	(0.278)	(0.270)	(0.290)	(0.282)	(0.281)	(0.282)
Log loan size part 3	1.232	0.875	1.219	0.875	1.219	0.875	1.206	0.875
	(0.163)	(0.328)	(0.161)	(0.328)	(0.166)	(0.338)	(0.164)	(0.338)
Log loan size part 4	2.691	0.971	2.722	0.972	2.739	0.950	2.768	0.951
	(0.255)	(0.260)	(0.258)	(0.261)	(0.256)	(0.252)	(0.258)	(0.252)
Loan age variables								
Loan age part 1	1.151	1.520	1.151	1.520	1.150	1.519	1.150	1.519
	(0.009)	(0.176)	(0.009)	(0.176)	(0.009)	(0.176)	(0.009)	(0.176)
Loan age part 2	1.003	1.080	1.003	1.080	1.003	1.081	1.003	1.081
	(0.004)	(0.018)	(0.004)	(0.018)	(0.004)	(0.018)	(0.004)	(0.018)
Loan age part 3	0.998	0.994	0.997	0.994	0.998	0.993	0.997	0.993
	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)
Loan age part 4	1.006	0.994	1.005	0.994	1.006	0.994	1.005	0.994
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables								
The credit score part 1	6.056	0.000	--	0.000	5.833	0.000	--	0.000
	(4.128)	(0.001)		(0.001)	(3.963)	(0.000)		(0.000)
The credit score part 2	31.343	0.003	--	0.003	24.393	0.002	--	0.002
	(29.425)	(0.007)		(0.006)	(22.369)	(0.004)		(0.004)
The credit score part 3	2.700	0.066	--	0.065	2.512	0.072	--	0.070
	(3.326)	(0.259)		(0.254)	(3.174)	(0.288)		(0.283)
The credit score part 4	1,011.740	0.000	--	0.000	901.847	0.000	--	0.000
	(1,681.227)	(0.000)		(0.000)	(1,502.466)	(0.000)		(0.000)
The debt-to-income ratio variables								
DTI part 1	0.419	3.171	0.334	3.155	0.422	3.269	0.341	3.253

	(0.157)	(4.951)	(0.125)	(4.926)	(0.158)	(5.107)	(0.127)	(5.083)
DTI part 2	2.578	16.107	1.957	16.038	2.687	18.995	2.063	18.916
	(1.298)	(25.450)	(0.985)	(25.341)	(1.354)	(30.050)	(1.039)	(29.925)
DTI part 3	0.514	2.104	0.435	2.097	0.571	2.454	0.486	2.446
	(0.282)	(3.046)	(0.238)	(3.035)	(0.313)	(3.562)	(0.266)	(3.549)
DTI part 4	1.377	1.911	1.507	1.915	1.328	1.986	1.457	1.990
	(0.481)	(1.456)	(0.525)	(1.458)	(0.463)	(1.511)	(0.507)	(1.515)
The dummy for the number of units	0.578	0.588	0.608	0.588	0.577	0.572	0.605	0.572
	(0.131)	(0.419)	(0.138)	(0.419)	(0.131)	(0.407)	(0.137)	(0.407)
Time period variables								
Time period part 1	1.023	0.996	1.024	0.996	1.023	0.997	1.023	0.997
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	0.963	1.036	0.964	1.036	0.964	1.034	0.965	1.034
	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	1.017	1.004	1.020	1.004	1.016	1.004	1.018	1.004
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)
	Number of mortgages 12,734				Number of mortgages 12,556			

The effect of the value of the call option supports H2. Because the range for any spline segment is less than one, it is meaningless to discuss a unit increase in the call option. Therefore, the odds ratios for the value of the call option are not discussed. Instead, this paper directly looks at the multinomial logit coefficients. This holds true throughout the paper whenever any explanatory variable has a spline segment less than one. When the value of the call option is below -0.025, the coefficient is 5.365 (the standard error is 0.884). When the value of the call option ranges between -0.025 and 0.037, the coefficient increases to 7.492 (the standard error is 0.892). When the value of the call option ranges between 0.037 and 0.102, the coefficient again increases to 7.927 (the standard error is 0.674). However, when the value of the call option is above 0.102, the coefficient drops to 1.892 (the standard error is 0.427). The significantly positive effect supports the argument that when the value of the call option is "in the money," households have more incentive to prepay their mortgages.

The effect of the monthly unemployment rate supports H3. When the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.312 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity increase by 1.117 times with a 1 percent increase in the unemployment rate.

The effect of negative equity is consistent with H4 and it supports the argument that when the put option is "in the money," households have more incentive to default their mortgages. When negative equity is below 18.366, the odds of default relative to continuity increase by 1.012 times with a 1 unit increase in negative equity. When negative equity ranges between 18.366 and 37.804, the odds of default relative to continuity increase by 1.019 times with a 1 unit increase in negative equity. When the range is between 37.804 and 62.910, the odds of default relative to continuity increase by 1.014 times with a 1 unit increase in negative equity.

However, when the negative equity is above 62.910, the odds of default relative to continuity will not change with a 1 unit increase.

Another important variable to evaluate the effect of negative equity is the negative equity dummy. The results support H5. The odds for mortgages with negative equity being defaulted instead of continued are 2.545 times as high as the same odds for mortgages with non-negative equity.

The results of original loan-to-value partially support H6. However, when its value ranges between 0.780 and 0.800, the effect becomes insignificantly negative. When original loan-to-value is below 0.650, the coefficient is 3.230 and the standard error is 1.172. In ranges between 0.650 and 0.780, the coefficient drops to 1.738 and the standard error is 1.227. When value ranges are between 0.780 and 0.800, the coefficient becomes insignificantly negative, which is -6.256 and the standard error is 5.353. In ranges between 0.800 and 0.950, the coefficient becomes significantly positive again, which is 3.369 and the standard error is 0.754. Finally, when original loan-to-value is above 0.950, the coefficient drops to 2.420 and the standard error is 4.562.

Consistent with the results in previous studies, H7 is held with the results in this paper. Moreover, the results of this study indicate the credit score has a strongly positive effect on the prepayment decision. The coefficient is 1.801 for prepayment (the standard error is 0.376) and is -7.953 for default (the standard error is 1.481) when the credit score is below 0.688. The coefficient increases to 3.445 for prepayment (the standard error is 0.939) and to -5.924 for default (the standard error is 2.484) when the credit score ranges between 0.688 and 0.736. However, the coefficient decreases to 0.993 for prepayment (the standard error is 1.232) and increases to -2.714 for default (the standard error is 3.907) when the credit score ranges between

0.736 and 0.774. The coefficient dramatically increases to 6.919 for prepayment (the standard error is 1.662) and drops to -20.685 for default (the standard error is 6.741) when the credit score is above 0.744. These results indicate both termination risks are sensitive to the high-level credit score. The households with high credit scores are more likely to prepay and less likely to default.

The effect of the debt-to-income ratio is consistent with H8. However, for the prepayment risk, the effect of the debt-to-income ratio is inconsistent and its effect on default is insignificant. When the debt-to-income ratio ranges below 0.250 and between 0.340 and 0.420, its effect on the prepayment risk is insignificantly negative. When the debt-to-income ratio ranges between 0.250 and 0.340 and above 0.420, its effect on the prepayment risk becomes insignificantly positive. For the default risk, when the debt-to-income ratio is below 0.250, the coefficient is 1.154 and the standard error is 1.561. When the ratio ranges between 0.250 and 0.340, the coefficient increases to 2.779 and the standard error is 1.580. When the debt-to-income ratio ranges between 0.340 and 0.420, the coefficient drops to 0.774 and the standard error is 1.448, and the coefficient further decreases to 0.648 and the standard error is 0.762 when the debt-to-income ratio ranges above 0.420.

The dummy for the number of units has a negative effect on both prepayment and default, and the effect is significant for prepayment. The odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.578 times as high as the same odds for mortgages covering only one house unit. Moreover, the odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.588 times as high as the same odds for mortgages covering only one house unit. The results show the mortgages for larger houses are less likely to be prepaid or defaulted than the mortgages for smaller houses.

Log loan size has a significantly positive effect on the prepayment; however, the effect is insignificant when the log loan size ranges between $149,941 and $210,029. On the other hand, its effect on the default is insignificant. The odds of prepayment relative to continuity increase by 2.128 times with a 1 unit increase in log loan size when it is below $105,030. In this range, the odds of default relative to continuity increase by 1.418 times with a 1 unit increase in log loan size. When ranges are between $105,030 and $149,941, the odds of prepayment relative to continuity increase by 2.317 times with a 1 unit increase in log loan size. In this range, the odds of default relative to continuity decrease by 0.632 times with a 1 unit increase in log loan size. However, the odds of prepayment relative to continuity increase by 1.232 times with a 1 unit increase in log loan size when ranges are between $149,941 and $210,029. In this range, the odds of default relative to continuity decrease by 0.875 times with a 1 unit increase in log loan size. Furthermore, when size is above $210,029, the odds of prepayment relative to continuity increase by 2.691 times with a 1 unit increase in log loan size. In this range, the odds of default relative to continuity decrease by 0.971 times with a 1 unit increase in log loan size.

Generally, loan age has a strongly positive relationship with the prepayment, but its effect is insignificantly negative when its value is between 24 and 45 months. When the loan age is below 24 months, the effect of loan age on the default decision is significantly positive. However, when loan age is above 24 months, its effect on default becomes significantly negative. When loan age is below 11 months, the odds of prepayment relative to continuity increase by 1.151 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity increase by 1.520 times with a 1 unit increase in loan age. When ranges are between 11 and 24 months, the odds of prepayment relative to continuity increase by 1.003 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity increase by 1.080

times with a 1 unit increase in loan age. When loan age ranges between 24 and 45 months, the odds of prepayment relative to continuity rapidly decrease by 0.998 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity decrease by 0.994 times with a 1 unit increase in loan age. When loan age is above 45 months, the odds of prepayment relative to continuity increase again by 1.006 times with a 1 unit increase in loan age, and the odds ratio for default remains the same as 0.994 in this range.

The time period in each model is also controlled. The results show mortgages are more likely to be prepaid from March 1999 to July 2003 and are less likely to be prepaid from July 2003 to November 2008. Prepayment increases again after November 2008. On the other hand, before January 2008, mortgages are less likely to default. However, starting from January 2010, mortgages become more likely to default.

It is also important to note that the strongly negative effect of the credit score still holds in the model where the credit score explains only the default risk. However, comparing the multinomial logit coefficients of the credit score in this model with those in the previous model, results indicate the coefficients in this model are smaller than those in the previous model but the standard errors do not change. When the credit score is below 0.688, the coefficient is -7.980 and the standard error is 1.481. With ranges between 0.688 and 0.736, the coefficient increases to -5.986 and the standard error is 2.484. When the credit score ranges are between 0.736 and 0.774, the coefficient increases to -2.732 and the standard error is 3.907. However, when the credit score is above 0.744, the coefficient dramatically drops to -20.819 and the standard error is 6.741. This result indicates households with high credit scores are much less likely to default.

Comparing the results of other explanatory variables in this model with those in the previous model, findings indicate the effect of the unemployment rate, negative equity, negative equity dummy, original loan-to-value, loan age, and time period remains the same. The value of the prepayment penalty dummy supports H1, but the odds ratios slightly increase in this model. The effect of the value of the call option supports H2. However, the coefficients in this model are smaller than that in the previous model. H8 is not rejected in this model and the effect of the debt-to-income ratio is still not significant for both prepayment and default. The dummy for the number of units still has a negative effect on both the prepayment and the default, and its effect for prepayment is stronger than that in the previous model. Comparing the odds ratios of log loan size in this model with those in the previous model, findings indicate the effect of log loan size on the default risk does not change between these two models. When the log loan size is below $210029, the odds ratios in this model are smaller than those in the previous model. However, when log loan size is above $210029, the odds ratio in this model is larger than that in the previous model.

The second group of models includes negative equity and negative equity dummy as explanatory variables. Table 17 contains the coefficients and table 18 contains the odds ratios. The following explanations are based on the dataset with right censored mortgages and the results for the dataset with modified mortgages deleted, which can be found in tables 17 and 18. Comparing the results of this model with the results of the model using negative equity, negative equity dummy and original loan-to-value as explanatory variables, results indicate the significance and coefficients of all explanatory variables for the prepayment risk are the same. Therefore, the discussion below only includes the results for the default risk. The results of the model with the credit score for both termination risks are discussed first and then are compared with the results of the model with the credit score only for the default risk.

[insert Tables 17 and 18 are here]

Table 17. Coefficients for Phoenix: Prepayment and default are dependent variables, negative equity and the negative equity dummy are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the default		The credit score only for default		The credit score for both the prepayment and the default		The credit score only for default	
	Prepayment	Default	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.813	--	-0.774	--	-0.816	--	-0.781	--
	(0.232)		(0.232)		(0.232)		(0.232)	
Call option variables								
Call option part 1	5.365	--	5.348	--	5.418	--	5.415	--
	(0.884)		(0.883)		(0.885)		(0.884)	
Call option part 2	7.491	--	7.336	--	7.557	--	7.404	--
	(0.892)		(0.892)		(0.891)		(0.890)	
Call option part 3	7.929	--	7.543	--	8.001	--	7.639	--
	(0.674)		(0.673)		(0.676)		(0.675)	
Call option part 4	1.890	--	1.312	--	1.999	--	1.449	--
	(0.427)		(0.425)		(0.428)		(0.425)	
The unemployment rate variables								
The unemployment rate part 1	--	0.242	--	0.242	--	0.251	--	0.251
		(0.104)		(0.104)		(0.104)		(0.104)
The unemployment rate part 2	--	0.094	--	0.094	--	0.097	--	0.097
		(0.058)		(0.058)		(0.058)		(0.058)
Negative equity variables								
Negative equity part 1	--	0.014	--	0.014	--	0.014	--	0.014
		(0.014)		(0.014)		(0.014)		(0.014)
Negative equity part 2	--	0.023	--	0.023	--	0.024	--	0.024
		(0.009)		(0.009)		(0.009)		(0.009)
Negative equity part 3	--	0.015	--	0.015	--	0.015	--	0.015
		(0.005)		(0.005)		(0.005)		(0.005)
Negative equity part 4	--	0.004	--	0.004	--	0.005	--	0.005
		(0.003)		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	1.022	--	1.022	--	1.028	--	1.028
		(0.211)		(0.211)		(0.212)		(0.212)
Log loan size variables								
Log loan size part 1	0.755	0.587	0.667	0.587	0.756	0.582	0.671	0.582
	(0.079)	(0.366)	(0.078)	(0.366)	(0.079)	(0.366)	(0.078)	(0.366)
Log loan size part 2	0.840	-0.579	0.809	-0.578	0.846	-0.545	0.817	-0.544
	(0.124)	(0.425)	(0.124)	(0.425)	(0.124)	(0.427)	(0.124)	(0.427)
Log loan size part 3	0.209	-0.124	0.198	-0.124	0.198	-0.113	0.188	-0.113
	(0.132)	(0.377)	(0.132)	(0.377)	(0.136)	(0.388)	(0.136)	(0.388)
Log loan size part 4	0.990	-0.288	1.001	-0.286	1.008	-0.296	1.018	-0.295
	(0.095)	(0.269)	(0.095)	(0.269)	(0.093)	(0.265)	(0.093)	(0.265)
Loan age variables								
Loan age part 1	0.140	0.423	0.140	0.423	0.140	0.422	0.140	0.422
	(0.008)	(0.116)	(0.008)	(0.116)	(0.008)	(0.116)	(0.008)	(0.116)
Loan age part 2	0.003	0.074	0.003	0.074	0.003	0.074	0.003	0.074
	(0.004)	(0.017)	(0.004)	(0.017)	(0.004)	(0.017)	(0.004)	(0.017)
Loan age part 3	-0.002	-0.012	-0.003	-0.012	-0.002	-0.012	-0.003	-0.012
	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)
Loan age part 4	0.006	-0.003	0.005	-0.003	0.006	-0.003	0.005	-0.003
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score								

variables								
The credit score part 1	1.801	-7.481	--	-7.508	1.764	-7.743	--	-7.771
	(0.682)	(1.471)		(1.471)	(0.679)	(1.464)		(1.464)
The credit score part 2	3.445	-6.706	--	-6.768	3.194	-7.185	--	-7.244
	(0.939)	(2.475)		(2.475)	(0.917)	(2.419)		(2.419)
The credit score part 3	0.993	-2.566	--	-2.584	0.921	-2.431	--	-2.448
	(1.232)	(3.898)		(3.898)	(1.264)	(4.007)		(4.007)
The credit score part 4	6.919	-21.403	--	-21.534	6.804	-21.545	--	-21.675
	(1.662)	(6.728)		(6.728)	(1.666)	(6.750)		(6.750)
The debt-to-income ratio variables								
DTI part 1	-0.870	1.291	-1.097	1.286	-0.863	1.332	-1.077	1.327
	(0.376)	(1.566)	(0.374)	(1.566)	(0.376)	(1.568)	(0.374)	(1.568)
DTI part 2	0.947	2.996	0.671	2.992	0.989	3.124	0.724	3.119
	(0.504)	(1.582)	(0.503)	(1.582)	(0.504)	(1.584)	(0.503)	(1.584)
DTI part 3	-0.666	0.783	-0.832	0.779	-0.561	0.919	-0.722	0.915
	(0.548)	(1.447)	(0.547)	(1.447)	(0.548)	(1.449)	(0.547)	(1.449)
DTI part 4	0.320	0.781	0.410	0.783	0.284	0.859	0.377	0.861
	(0.349)	(0.757)	(0.348)	(0.757)	(0.349)	(0.756)	(0.348)	(0.756)
The dummy for the number of units	-0.548	-0.625	-0.498	-0.625	-0.551	-0.646	-0.503	-0.646
	(0.227)	(0.711)	(0.227)	(0.711)	(0.227)	(0.711)	(0.227)	(0.711)
Time period variables								
Time period part 1	0.023	-0.008	0.023	-0.008	0.022	-0.007	0.023	-0.007
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	-0.037	0.037	-0.037	0.037	-0.037	0.035	-0.036	0.035
	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	0.017	0.004	0.019	0.004	0.016	0.004	0.018	0.004
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.009)

Number of mortgages 12,734 Number of mortgages 12,556

Table 18. Odds ratios for Phoenix: Prepayment and default are dependent variables, negative equity and the negative equity dummy are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the default		The credit score only for default		The credit score for both the prepayment and the default		The credit score only for default	
	Prepayment	Default	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	0.444	--	0.461	--	0.442	--	0.458	--
	(0.103)		(0.107)		(0.102)		(0.106)	
Call option variables								
Call option part 1	213.802	--	210.244	--	225.445	--	224.642	--
	(189.055)		(185.745)		(199.439)		(198.555)	
Call option part 2	1,792.521	--	1,534.332	--	1,914.403	--	1,643.356	--
	(1,599.044)		(1,367.936)		(1,705.068)		(1,462.802)	
Call option part 3	2,775.324	--	1,888.352	--	2,984.801	--	2,078.085	--
	(1,871.113)		(1,270.746)		(2,017.162)		(1,401.864)	
Call option part 4	6.620	--	3.715	--	7.384	--	4.258	--
	(2.828)		(1.578)		(3.158)		(1.811)	
The unemployment rate variables								
The unemployment rate part 1	--	1.274	--	1.274	--	1.285	--	1.285
		(0.133)		(0.133)		(0.134)		(0.134)
The unemployment rate part 2	--	1.099	--	1.099	--	1.102	--	1.102
		(0.064)		(0.064)		(0.064)		(0.064)

Negative equity variables								
Negative equity part 1	--	1.014	--	1.014	--	1.014	--	1.014
		(0.014)		(0.014)		(0.014)		(0.014)
Negative equity part 2	--	1.024	--	1.024	--	1.024	--	1.024
		(0.009)		(0.009)		(0.009)		(0.009)
Negative equity part 3	--	1.015	--	1.015	--	1.015	--	1.015
		(0.006)		(0.006)		(0.006)		(0.006)
Negative equity part 4	--	1.004	--	1.004	--	1.005	--	1.005
		(0.003)		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	2.779	--	2.778	--	2.795	--	2.794
		(0.586)		(0.586)		(0.591)		(0.591)
Log loan size variables								
Log loan size part 1	2.128	1.799	1.949	1.798	2.130	1.790	1.956	1.789
	(0.168)	(0.658)	(0.152)	(0.658)	(0.168)	(0.655)	(0.153)	(0.655)
Log loan size part 2	2.317	0.560	2.246	0.561	2.330	0.580	2.264	0.580
	(0.287)	(0.238)	(0.278)	(0.239)	(0.290)	(0.248)	(0.281)	(0.248)
Log loan size part 3	1.232	0.884	1.219	0.884	1.219	0.893	1.206	0.893
	(0.163)	(0.333)	(0.161)	(0.333)	(0.166)	(0.347)	(0.164)	(0.347)
Log loan size part 4	2.692	0.750	2.722	0.751	2.739	0.743	2.768	0.745
	(0.255)	(0.201)	(0.258)	(0.202)	(0.256)	(0.197)	(0.258)	(0.198)
Loan age variables								
Loan age part 1	1.151	1.527	1.151	1.527	1.150	1.525	1.150	1.525
	(0.009)	(0.177)	(0.009)	(0.177)	(0.009)	(0.177)	(0.009)	(0.177)
Loan age part 2	1.003	1.076	1.003	1.076	1.003	1.077	1.003	1.077
	(0.004)	(0.018)	(0.004)	(0.018)	(0.004)	(0.018)	(0.004)	(0.018)
Loan age part 3	0.998	0.988	0.997	0.988	0.998	0.988	0.997	0.988
	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)
Loan age part 4	1.006	0.997	1.005	0.997	1.006	0.997	1.005	0.997
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables								
The credit score part 1	6.057	0.001	--	0.001	5.834	0.000	--	0.000
	(4.129)	(0.001)		(0.001)	(3.963)	(0.001)		(0.001)
The credit score part 2	31.347	0.001	--	0.001	24.394	0.001	--	0.001
	(29.429)	(0.003)		(0.003)	(22.371)	(0.002)		(0.002)
The credit score part 3	2.700	0.077	--	0.075	2.512	0.088	--	0.086
	(3.326)	(0.299)		(0.294)	(3.174)	(0.352)		(0.346)
The credit score part 4	1,011.395	0.000	--	0.000	901.466	0.000	--	0.000
	(1,680.652)	(0.000)		(0.000)	(1,501.828)	(0.000)		(0.000)
The debt-to-income ratio variables								
DTI part 1	0.419	3.637	0.334	3.618	0.422	3.787	0.341	3.769
	(0.157)	(5.697)	(0.125)	(5.668)	(0.158)	(5.937)	(0.127)	(5.908)
DTI part 2	2.578	20.010	1.957	19.924	2.688	22.728	2.063	22.633
	(1.298)	(31.665)	(0.985)	(31.528)	(1.354)	(36.003)	(1.039)	(35.851)
DTI part 3	0.514	2.188	0.435	2.180	0.571	2.506	0.486	2.497
	(0.282)	(3.166)	(0.238)	(3.155)	(0.313)	(3.632)	(0.266)	(3.619)
DTI part 4	1.377	2.183	1.507	2.188	1.328	2.361	1.457	2.366
	(0.481)	(1.654)	(0.525)	(1.657)	(0.463)	(1.785)	(0.507)	(1.789)
The dummy for the number of units	0.578	0.535	0.608	0.535	0.577	0.524	0.605	0.524
	(0.131)	(0.381)	(0.138)	(0.381)	(0.131)	(0.373)	(0.137)	(0.373)
Time period variables								
Time period part 1	1.023	0.992	1.024	0.992	1.023	0.993	1.023	0.993
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	0.963	1.038	0.964	1.038	0.964	1.036	0.965	1.036
	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	1.017	1.004	1.020	1.004	1.016	1.004	1.018	1.004
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.009)

Number of mortgages 12,734 Number of mortgages 12,556

The third column in table 17 shows that the effect of the monthly unemployment rate supports H3. When the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.274 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity will increase by 1.099 times with a 1 percent increase in the unemployment rate.

The effect of negative equity is consistent with H4, which supports the argument that when the put option is “in the money,” households have more incentive to default their mortgages. When negative equity is below 18.366, the odds of default relative to continuity increase by 1.014 times with a 1 unit increase in negative equity, and when its value ranges between 18.366 and 37.804, the odds of default relative to continuity increase by 1.024 times with a 1 unit increase in negative equity. When the negative equity ranges between 37.804 and 62.910, the odds of default relative to continuity increase by 1.015 times with a 1 unit increase in negative equity, and when its value is above 62.910, the odds of default relative to continuity increase by 1.004 times with a 1 unit increase in negative equity.

The results of the negative equity dummy support H5. The odds for mortgages with negative equity being defaulted instead of continued are 2.779 times as high as the same odds for mortgages with non-negative equity.

The results of the credit score support H7. The coefficient is -7.481 (the standard error is 1.471) when the credit score is below 0.688 and the coefficient increases to -6.706 (the standard error is 2.475) when the credit score ranges between 0.688 and 0.736. The coefficient further increases to -2.566 (the standard error is 3.898) when the score ranges between 0.736 and 0.774 and the coefficient dramatically drops to -21.403 (the standard error is 6.728) when the credit score is above 0.744. These results indicate households with high credit scores are much less likely to default.

The effect of the debt-to-income ratio on the default risk supports H8 but the results are not significant. The dummy for the number of units has a negative effect on the default risk. The odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.535 times as high as the same odds for mortgages covering only one house unit. These results indicate mortgages for larger houses are less likely to default.

Log loan size has an insignificantly negative effect on the default. However, when log loan size is below $105,030, its effect becomes insignificantly positive and the odds of default relative to continuity increase by 1.799 times with a 1 unit increase in log loan size. When size ranges between $105,030 and $149,941, the odds of default relative to continuity decrease by 0.560 times with a 1 unit increase in log loan size, and when its value ranges between $149,941 and $210,029, the odds of default relative to continuity decrease by 0.884 times with a 1 unit increase in log loan size. When log loan size is above $210,029, the odds of default relative to continuity decrease by 0.750 times with a 1 unit increase in log loan size.

Loan age has a significantly positive effect on the default risk when the age is below 24 months, and its effect becomes significantly negative when the age is above 24 months. When loan age is below 11 months, the odds of default relative to continuity increase by 1.527 times with a 1 unit increase in loan age. When loan age ranges between 11 and 24 months, the odds of default relative to continuity increase by 1.076 times with a 1 unit increase in loan age. When ranges are between 24 and 45 months, the odds of default relative to continuity decrease by 0.988 times with a 1 unit increase in loan age, and when the age is above 45 months, the odds of default relative to continuity decrease by 0.997 times with a 1 unit increase in loan age.

H7 is still held with the results in the model using the credit score to explain only the default risk. Comparing the coefficients of the credit score for the default risk in this model with those in the model with the credit score used for both termination risks, findings indicate the coefficients in this model are smaller than those in the previous model; however, the standard errors do not change. When the credit score is below 0.688, the coefficient is -7.508 and the standard error is 1.471. When the range is between 0.688 and 0.736, the coefficient increases to -6.768 and the standard error is 2.475. When the credit score ranges between 0.736 and 0.774, the coefficient increases to -2.584 and the standard error is 3.898. However, when the credit score is above 0.744, the coefficient dramatically drops to -21.534 and the standard error is 6.728. This result indicates households with high credit scores are much less likely to default.

The results of other explanatory variables are compared with those in the model with the credit score used for both termination risks. The findings indicate the effect of the unemployment rate, negative equity, negative equity dummy, log loan size, loan age, unit dummy and time period remains the same. However, the coefficients of the debt-to-income ratio slightly decrease.

The third group of models includes the negative equity dummy and original loan-to-value as explanatory variables. Table 19 contains the coefficients and table 20 contains the odds ratios. The following explanations are based on the dataset with right censored mortgages and the results for the dataset with modified mortgages deleted can be found in tables 19 and 20. The results of this model are compared with the results of the model using negative equity, negative equity dummy and original loan-to-value as explanatory variables, and findings indicate the significance and the coefficients of all explanatory variables for the prepayment risk are the same. Therefore, only the results for the default risk are discussed. The results of the model with the credit score for both termination risks is first discussed and then these results are compared with the results of the model with the credit score only for the default risk.

[insert Tables 19 and 20 are here]

The effect of the monthly unemployment rate supports H3. When the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.351 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity increase by 1.155 times with a 1 percent increase in the unemployment rate.

The effect of the negative equity dummy supports H5. The odds for mortgages with negative equity being defaulted instead of continued are 3.947 times as high as the same odds for mortgages with non-negative equity.

H6 is partially supported in this model. When original loan-to-value is below 0.650, the coefficient is 3.213 (the standard error is 1.204), and when it ranges between 0.650 and 0.780, the coefficient increases to 3.465 (the standard error is 1.195). However, when original loan-to-value ranges between 0.780 and 0.800, the coefficient becomes insignificantly negative, which is -7.657 (the standard error is 5.314) and in ranges between 0.800 and 0.950, the coefficient becomes significantly positive again, which is 3.457 (the standard error is 0.747). When original loan-to-value is above 0.950, the coefficient increases to 5.414 and the standard error is 4.562. Furthermore, when original loan-to-value is kept in the model and negative equity is deleted, the absolute value of the coefficients of original loan-to-value increases.

The effect of the credit score supports H7. The coefficient is -7.904 (the standard error is 1.467) when the credit score is below 0.688 and the coefficient increases to -6.151 (the standard error is 2.475) when the score ranges between 0.688 and 0.736. The coefficient further increases

Table 19. Coefficients for Phoenix: Prepayment and default are dependent variables, the negative equity dummy and original loan-to-value are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the default		The credit score only for default		The credit score for both the prepayment and the default		The credit score only for default	
	Prepayment	Default	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.813	--	-0.774	--	-0.816	--	-0.781	--
	(0.232)		(0.232)		(0.232)		(0.232)	
Call option variables								
Call option part 1	5.364	--	5.348	--	5.417	--	5.414	--
	(0.884)		(0.883)		(0.885)		(0.884)	
Call option part 2	7.494	--	7.338	--	7.560	--	7.407	--
	(0.892)		(0.892)		(0.891)		(0.890)	
Call option part 3	7.924	--	7.538	--	7.996	--	7.634	--
	(0.674)		(0.673)		(0.676)		(0.675)	
Call option part 4	1.892	--	1.315	--	2.001	--	1.451	--
	(0.427)		(0.425)		(0.428)		(0.425)	
The unemployment rate variables								
The unemployment rate part 1	--	0.301	--	0.301	--	0.307	--	0.308
		(0.104)		(0.103)		(0.104)		(0.104)
The unemployment rate part 2	--	0.144	--	0.144	--	0.147	--	0.147
		(0.057)		(0.057)		(0.057)		(0.057)
The negative equity dummy	--	1.373	--	1.373	--	1.385	--	1.384
		(0.105)		(0.105)		(0.105)		(0.105)
Original loan-to-value variables								
Original LTV part 1	--	3.213	--	3.213	--	3.281	--	3.281
		(1.204)		(1.204)		(1.209)		(1.209)
Original LTV part 2	--	3.465	--	3.465	--	3.444	--	3.444
		(1.195)		(1.195)		(1.196)		(1.196)
Original LTV part 3	--	-7.657	--	-7.657	--	-9.227	--	-9.228
		(5.314)		(5.314)		(5.316)		(5.316)
Original LTV part 4	--	3.457	--	3.457	--	3.701	--	3.701
		(0.747)		(0.747)		(0.748)		(0.748)
Original LTV part 5	--	5.414	--	5.409	--	5.807	--	5.800
		(4.536)		(4.536)		(4.553)		(4.553)
Log loan size variables								
Log loan size part 1	0.755	0.241	0.667	0.241	0.756	0.238	0.671	0.237
	(0.079)	(0.374)	(0.078)	(0.374)	(0.079)	(0.374)	(0.078)	(0.374)
Log loan size part 2	0.840	-0.158	0.809	-0.157	0.846	-0.121	0.817	-0.121
	(0.124)	(0.421)	(0.124)	(0.421)	(0.124)	(0.423)	(0.124)	(0.423)
Log loan size part 3	0.209	0.540	0.198	0.540	0.198	0.570	0.188	0.570
	(0.132)	(0.346)	(0.132)	(0.346)	(0.136)	(0.357)	(0.136)	(0.357)
Log loan size part 4	0.990	0.095	1.001	0.096	1.008	0.147	1.018	0.149
	(0.095)	(0.222)	(0.095)	(0.222)	(0.093)	(0.219)	(0.093)	(0.219)
Loan age variables								
Loan age part 1	0.140	0.417	0.140	0.417	0.140	0.416	0.140	0.417
	(0.008)	(0.116)	(0.008)	(0.116)	(0.008)	(0.116)	(0.008)	(0.116)
Loan age part 2	0.003	0.083	0.003	0.083	0.003	0.084	0.003	0.084
	(0.004)	(0.017)	(0.004)	(0.017)	(0.004)	(0.017)	(0.004)	(0.017)
Loan age part 3	-0.002	0.001	-0.003	0.001	-0.002	0.000	-0.003	0.000
	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)
Loan age part 4	0.006	-0.009	0.005	-0.009	0.006	-0.009	0.005	-0.009
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score								

variables								
The credit score part 1	1.801	-7.904	--	-7.932	1.763	-8.099	--	-8.126
	(0.682)	(1.467)		(1.467)	(0.679)	(1.460)		(1.460)
The credit score part 2	3.445	-6.151	--	-6.212	3.194	-6.586	--	-6.645
	(0.939)	(2.475)		(2.475)	(0.917)	(2.420)		(2.420)
The credit score part 3	0.993	-3.126	--	-3.143	0.921	-3.104	--	-3.121
	(1.232)	(3.913)		(3.912)	(1.264)	(4.020)		(4.020)
The credit score part 4	6.919	-20.290	--	-20.425	6.804	-20.314	--	-20.448
	(1.662)	(6.783)		(6.783)	(1.666)	(6.804)		(6.804)
The debt-to-income ratio variables								
DTI part 1	-0.870	1.033	-1.097	1.028	-0.863	1.051	-1.077	1.046
	(0.376)	(1.571)	(0.374)	(1.571)	(0.376)	(1.572)	(0.374)	(1.572)
DTI part 2	0.947	3.232	0.671	3.228	0.989	3.444	0.724	3.440
	(0.504)	(1.582)	(0.503)	(1.582)	(0.504)	(1.584)	(0.503)	(1.584)
DTI part 3	-0.665	0.565	-0.832	0.561	-0.560	0.677	-0.722	0.674
	(0.548)	(1.447)	(0.547)	(1.447)	(0.548)	(1.450)	(0.547)	(1.450)
DTI part 4	0.320	0.653	0.410	0.655	0.284	0.692	0.377	0.694
	(0.349)	(0.761)	(0.348)	(0.761)	(0.349)	(0.761)	(0.348)	(0.761)
The dummy for the number of units	-0.548	-0.461	-0.498	-0.461	-0.551	-0.486	-0.503	-0.486
	(0.227)	(0.711)	(0.227)	(0.711)	(0.227)	(0.711)	(0.227)	(0.711)
Time period variables								
Time period part 1	0.023	-0.005	0.023	-0.005	0.022	-0.004	0.023	-0.004
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	-0.037	0.033	-0.037	0.033	-0.037	0.031	-0.036	0.031
	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	0.017	0.005	0.019	0.005	0.016	0.005	0.018	0.005
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)

Number of mortgages 12,734 Number of mortgages 12,556

Table 20. Odds ratios for Phoenix: Prepayment and default are dependent variables, the negative equity dummy and original loan-to-value are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the default		The credit score only for default		The credit score for both the prepayment and the default		The credit score only for default	
	Prepayment	Default	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	0.444	--	0.461	--	0.442	--	0.458	--
	(0.103)		(0.107)		(0.102)		(0.106)	
Call option variables								
Call option part 1	213.659	--	210.101	--	225.273	--	224.466	--
	(188.926)		(185.616)		(199.284)		(198.397)	
Call option part 2	1,796.933	--	1,538.343	--	1,919.442	--	1,647.935	--
	(1,602.983)		(1,371.515)		(1,709.560)		(1,466.880)	
Call option part 3	2,761.841	--	1,878.719	--	2,968.970	--	2,066.528	--
	(1,862.038)		(1,264.273)		(2,006.480)		(1,394.078)	
Call option part 4	6.634	--	3.723	--	7.399	--	4.267	--
	(2.834)		(1.582)		(3.164)		(1.815)	
The unemployment rate variables								
The unemployment rate part 1	--	1.351	--	1.351	--	1.360	--	1.360
		(0.140)		(0.140)		(0.141)		(0.141)
The unemployment rate part 2	--	1.155	--	1.155	--	1.159	--	1.159
		(0.066)		(0.066)		(0.066)		(0.066)
The negative equity dummy	--	3.947	--	3.946	--	3.993	--	3.992

		(0.415)		(0.415)		(0.420)		(0.420)
Original loan-to-value variables								
Original LTV part 1	--	24.859	--	24.849	--	26.607	--	26.597
		(29.924)		(29.913)		(32.160)		(32.149)
Original LTV part 2	--	31.965	--	31.970	--	31.320	--	31.327
		(38.205)		(38.212)		(37.461)		(37.469)
Original LTV part 3	--	0.000	--	0.000	--	0.000	--	0.000
		(0.003)		(0.003)		(0.001)		(0.001)
Original LTV part 4	--	31.709	--	31.719	--	40.469	--	40.480
		(23.674)		(23.681)		(30.278)		(30.286)
Original LTV part 5	--	224.510	--	223.355	--	332.720	--	330.394
		(1,018.460)		(1,013.212)		(1,514.944)		(1,504.337)
Log loan size variables								
Log loan size part 1	2.128	1.273	1.949	1.273	2.130	1.268	1.956	1.268
	(0.168)	(0.476)	(0.152)	(0.475)	(0.168)	(0.475)	(0.153)	(0.475)
Log loan size part 2	2.317	0.854	2.246	0.854	2.330	0.886	2.264	0.886
	(0.287)	(0.359)	(0.278)	(0.360)	(0.290)	(0.375)	(0.281)	(0.375)
Log loan size part 3	1.232	1.716	1.219	1.716	1.219	1.768	1.206	1.768
	(0.163)	(0.594)	(0.161)	(0.594)	(0.166)	(0.632)	(0.164)	(0.632)
Log loan size part 4	2.691	1.099	2.722	1.101	2.739	1.159	2.768	1.160
	(0.255)	(0.244)	(0.258)	(0.244)	(0.256)	(0.254)	(0.258)	(0.255)
Loan age variables								
Loan age part 1	1.151	1.518	1.151	1.518	1.150	1.517	1.150	1.517
	(0.009)	(0.177)	(0.009)	(0.177)	(0.009)	(0.176)	(0.009)	(0.177)
Loan age part 2	1.003	1.087	1.003	1.087	1.003	1.088	1.003	1.088
	(0.004)	(0.018)	(0.004)	(0.018)	(0.004)	(0.018)	(0.004)	(0.018)
Loan age part 3	0.998	1.001	0.997	1.001	0.998	1.000	0.997	1.000
	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)
Loan age part 4	1.006	0.991	1.005	0.991	1.006	0.991	1.005	0.991
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables								
The credit score part 1	6.054	0.000	--	0.000	5.831	0.000	--	0.000
	(4.127)	(0.001)		(0.001)	(3.961)	(0.000)		(0.000)
The credit score part 2	31.335	0.002	--	0.002	24.386	0.001	--	0.001
	(29.418)	(0.005)		(0.005)	(22.363)	(0.003)		(0.003)
The credit score part 3	2.701	0.044	--	0.043	2.512	0.045	--	0.044
	(3.327)	(0.172)		(0.169)	(3.175)	(0.180)		(0.177)
The credit score part 4	1,011.621	0.000	--	0.000	901.739	0.000	--	0.000
	(1,681.017)	(0.000)		(0.000)	(1,502.275)	(0.000)		(0.000)
The debt-to-income ratio variables								
DTI part 1	0.419	2.810	0.334	2.796	0.422	2.859	0.341	2.846
	(0.157)	(4.414)	(0.125)	(4.392)	(0.158)	(4.495)	(0.127)	(4.474)
DTI part 2	2.578	25.328	1.957	25.220	2.687	31.318	2.063	31.186
	(1.298)	(40.066)	(0.985)	(39.893)	(1.354)	(49.604)	(1.039)	(49.395)
DTI part 3	0.514	1.759	0.435	1.753	0.571	1.969	0.486	1.962
	(0.282)	(2.545)	(0.238)	(2.536)	(0.313)	(2.854)	(0.266)	(2.844)
DTI part 4	1.377	1.922	1.507	1.926	1.328	1.998	1.457	2.003
	(0.481)	(1.463)	(0.525)	(1.466)	(0.463)	(1.520)	(0.507)	(1.523)
The dummy for the number of units	0.578	0.631	0.608	0.631	0.577	0.615	0.605	0.615
	(0.131)	(0.449)	(0.138)	(0.449)	(0.131)	(0.438)	(0.137)	(0.438)
Time period variables								
Time period part 1	1.023	0.995	1.024	0.995	1.023	0.996	1.023	0.996
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	0.963	1.034	0.964	1.034	0.964	1.032	0.965	1.032
	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	1.017	1.005	1.020	1.005	1.016	1.005	1.018	1.005
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)
	Number of mortgages 12,734				Number of mortgages 12,556			

to -3.126 (the standard error is 3.913) when the credit score ranges between 0.736 and 0.774 and it drops dramatically to -20.290 (the standard error is 6.783) when the score is above 0.744. These results indicate households with high credit scores are much less likely to default.

The effect of the debt-to-income ratio on the default risk is consistent with H8 but the results are not significant. The dummy for the number of units has a negative effect on the default risk. The odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.631 times as high as the same odds for mortgages covering only one house unit. The results show mortgages for larger houses are less likely to default.

Different from the previous results, the effect of log loan size on the default risk becomes insignificant. The difference of the results among models indicates the effect of the log loan size is sensitive to the combination of the explanatory variables chosen.

Loan age has a significantly positive effect on the default risk when the age is below 45 months, and its effect becomes significantly negative when the age is above 45 months. When loan age is below 11 months, the odds of default relative to continuity increase by 1.518 times with a 1 unit increase in loan age. When the age ranges between 11 and 24 months, the odds of default relative to continuity increase by 1.087 times with a 1 unit increase in loan age, and when the value ranges between 24 and 45 months, the odds of default relative to continuity increase by 1.001 times with a 1 unit increase in loan age. However, when loan age is above 45 months, the odds of default relative to continuity decrease by 0.991 times with a 1 unit increase in loan age.

H7 is still held with the results in the model using the credit score to explain only the default risk. Comparing the coefficients of the credit score in this model with those in the model with the credit score used for both termination risks, findings indicate the coefficients in this model are smaller than those in the previous model. However, the standard errors do not change. When the credit score is below 0.688, the coefficient is -7.932 (the standard error is 1.467), and when the score ranges between 0.688 and 0.736, the coefficient increases to -6.212 (the standard error is 2.475). When the credit score ranges between 0.736 and 0.774, the coefficient increases to -3.143 (the standard error is 3.912), and when the credit score is above 0.744, the coefficient dramatically drops to -20.425 (the standard error is 6.783). This result indicates households with high credit scores are much less likely to default.

The results of other explanatory variables for the default risk in this model are compared with those in the model with the credit score only used for both termination risks, and findings indicate the effect of the unemployment rate, negative equity dummy, original loan-to-value, log loan size, loan age, unit dummy and time period remains the same. However, the coefficients of the debt-to-income ratio slightly decrease.

Group 2: Prepayment and 90-days-delinquency are dependent variables

The first group of models includes negative equity and the negative equity dummy as explanatory variables. The coefficients are listed in table 21 and the odds ratios are listed in table 22. The following explanations are based on the dataset with right censored mortgages. The results for the dataset with modified mortgages deleted can be found in tables 21 and 22. The results of the model with credit score for both termination risks is first introduced followed by a comparison of the results of the model with the credit score only for the 90-days-delinquency risk.

[insert Tables 21 and 22 are here]

Table 21. Coefficients for Phoenix: Prepayment and 90-days-delinquency are dependent variables, negative equity and the negative equity dummy are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency		The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	-0.819	--	-0.785	--	-0.823	--	-0.792	--
	(0.232)		(0.232)		(0.232)		(0.232)	
Call option variables								
Call option part 1	5.402	--	5.391	--	5.446	--	5.448	--
	(0.893)		(0.892)		(0.893)		(0.892)	
Call option part 2	7.349	--	7.214	--	7.424	--	7.291	--
	(0.911)		(0.910)		(0.910)		(0.910)	
Call option part 3	8.383	--	8.043	--	8.438	--	8.115	--
	(0.683)		(0.682)		(0.684)		(0.683)	
Call option part 4	2.177	--	1.668	--	2.285	--	1.798	--
	(0.424)		(0.422)		(0.424)		(0.423)	
The unemployment rate variables								
The unemployment rate part 1	--	0.295	--	0.295	--	0.288	--	0.289
		(0.076)		(0.076)		(0.079)		(0.079)
The unemployment rate part 2	--	0.177	--	0.177	--	0.187	--	0.187
		(0.051)		(0.051)		(0.054)		(0.054)
Negative equity variables								
Negative equity part 1	--	0.014	--	0.014	--	0.019	--	0.019
		(0.014)		(0.014)		(0.014)		(0.014)
Negative equity part 2	--	0.022	--	0.022	--	0.023	--	0.023
		(0.008)		(0.008)		(0.009)		(0.009)
Negative equity part 3	--	0.018	--	0.018	--	0.018	--	0.018
		(0.005)		(0.005)		(0.005)		(0.005)
Negative equity part 4	--	0.003	--	0.003	--	0.004	--	0.004
		(0.003)		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	0.849	--	0.849	--	0.859	--	0.858
		(0.197)		(0.197)		(0.211)		(0.211)
Log loan size variables								
Log loan size part 1	0.786	0.325	0.704	0.325	0.787	0.368	0.707	0.368
	(0.079)	(0.308)	(0.079)	(0.308)	(0.080)	(0.324)	(0.080)	(0.324)
Log loan size part 2	0.836	-0.041	0.808	-0.040	0.838	-0.294	0.811	-0.293
	(0.125)	(0.383)	(0.124)	(0.383)	(0.121)	(0.391)	(0.121)	(0.391)
Log loan size part 3	0.219	-0.552	0.209	-0.552	0.218	-0.453	0.208	-0.453
	(0.135)	(0.347)	(0.135)	(0.347)	(0.136)	(0.370)	(0.136)	(0.370)
Log loan size part 4	0.982	-0.229	0.990	-0.227	1.004	-0.279	1.012	-0.277
	(0.094)	(0.246)	(0.094)	(0.246)	(0.094)	(0.259)	(0.093)	(0.259)
Loan age variables								
Loan age part 1	0.140	0.244	0.140	0.244	0.139	0.232	0.139	0.232
	(0.008)	(0.048)	(0.008)	(0.048)	(0.008)	(0.048)	(0.008)	(0.048)
Loan age part 2	0.003	0.037	0.003	0.037	0.003	0.033	0.003	0.033
	(0.004)	(0.014)	(0.004)	(0.014)	(0.004)	(0.014)	(0.004)	(0.014)
Loan age part 3	-0.002	-0.010	-0.003	-0.010	-0.002	-0.010	-0.003	-0.010
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	0.006	-0.003	0.005	-0.003	0.006	-0.003	0.005	-0.003
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)

The credit score variables								
The credit score part 1	1.347	-8.144	--	-8.166	1.338	-8.264	--	-8.287
	(0.675)	(1.270)		(1.270)	(0.665)	(1.329)		(1.329)
The credit score part 2	3.455	-9.572	--	-9.632	3.227	-9.459	--	-9.516
	(0.939)	(2.291)		(2.291)	(0.937)	(2.409)		(2.409)
The credit score part 3	0.983	-5.221	--	-5.238	0.946	-3.513	--	-3.530
	(1.269)	(3.874)		(3.874)	(1.306)	(4.134)		(4.134)
The credit score part 4	6.571	-21.612	--	-21.727	6.456	-21.387	--	-21.502
	(1.668)	(6.631)		(6.631)	(1.673)	(6.778)		(6.778)
The debt-to-income ratio variables								
DTI part 1	-0.884	2.441	-1.091	2.436	-0.886	2.173	-1.074	2.168
	(0.376)	(1.536)	(0.374)	(1.536)	(0.380)	(1.570)	(0.378)	(1.570)
DTI part 2	1.038	3.672	0.785	3.668	1.159	2.972	0.884	2.968
	(0.505)	(1.474)	(0.505)	(1.474)	(0.567)	(1.738)	(0.566)	(1.738)
DTI part 3	-0.679	1.063	-0.851	1.059	-0.424	1.850	-0.576	1.846
	(0.550)	(1.325)	(0.549)	(1.325)	(0.486)	(1.241)	(0.485)	(1.241)
DTI part 4	0.354	0.675	0.435	0.678	0.298	0.596	0.378	0.598
	(0.351)	(0.695)	(0.350)	(0.695)	(0.347)	(0.728)	(0.346)	(0.728)
The dummy for the number of units	-0.610	-0.087	-0.567	-0.087	-0.613	-0.298	-0.573	-0.299
	(0.233)	(0.505)	(0.232)	(0.505)	(0.233)	(0.582)	(0.232)	(0.582)
Time period variables								
Time period part 1	0.022	-0.001	0.023	-0.001	0.022	-0.001	0.023	-0.001
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	-0.037	0.003	-0.036	0.003	-0.036	-0.003	-0.035	-0.003
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	0.016	0.007	0.019	0.007	0.015	0.011	0.018	0.011
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.009)	(0.001)	(0.009)

Number of mortgages 12,698 Number of mortgages 12,522

Table 22. Odds ratios for Phoenix: Prepayment and 90-days-delinquency are dependent variables, negative equity and the negative equity dummy are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency		The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	0.441	--	0.456	--	0.439	--	0.453	--
	(0.102)		(0.106)		(0.102)		(0.105)	
Call option variables								
Call option part 1	221.743	--	219.452	--	231.808	--	232.186	--
	(197.936)		(195.743)		(206.889)		(207.069)	
Call option part 2	1,555.246	--	1,358.696	--	1,675.123	--	1,466.666	--
	(1,416.409)		(1,236.804)		(1,525.158)		(1,334.700)	
Call option part 3	4,373.069	--	3,111.915	--	4,618.155	--	3,345.874	--
	(2,985.586)		(2,121.109)		(3,158.802)		(2,284.934)	
Call option part 4	8.823	--	5.299	--	9.829	--	6.039	--
	(3.740)		(2.237)		(4.172)		(2.553)	
The unemployment rate variables								
The unemployment rate part 1	--	1.343	--	1.344	--	1.334	--	1.335
		(0.102)		(0.102)		(0.106)		(0.106)
The unemployment rate part 2	--	1.194	--	1.194	--	1.206	--	1.206

		(0.061)		(0.061)		(0.065)		(0.065)
Negative equity variables								
Negative equity part 1	--	1.014	--	1.014	--	1.020	--	1.020
		(0.014)		(0.014)		(0.015)		(0.015)
Negative equity part 2	--	1.023	--	1.023	--	1.023	--	1.023
		(0.009)		(0.009)		(0.009)		(0.009)
Negative equity part 3	--	1.018	--	1.018	--	1.018	--	1.018
		(0.005)		(0.005)		(0.006)		(0.006)
Negative equity part 4	--	1.003	--	1.003	--	1.004	--	1.004
		(0.003)		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	2.337	--	2.337	--	2.360	--	2.359
		(0.461)		(0.461)		(0.497)		(0.497)
Log loan size variables								
Log loan size part 1	2.195	1.384	2.022	1.383	2.196	1.445	2.028	1.445
	(0.174)	(0.427)	(0.159)	(0.427)	(0.176)	(0.468)	(0.161)	(0.468)
Log loan size part 2	2.308	0.960	2.245	0.961	2.311	0.745	2.250	0.746
	(0.287)	(0.368)	(0.279)	(0.368)	(0.281)	(0.291)	(0.273)	(0.291)
Log loan size part 3	1.245	0.576	1.232	0.576	1.244	0.636	1.231	0.636
	(0.168)	(0.200)	(0.166)	(0.200)	(0.169)	(0.235)	(0.168)	(0.236)
Log loan size part 4	2.670	0.795	2.691	0.797	2.730	0.757	2.752	0.758
	(0.252)	(0.196)	(0.253)	(0.196)	(0.255)	(0.196)	(0.257)	(0.196)
Loan age variables								
Loan age part 1	1.150	1.276	1.150	1.276	1.149	1.261	1.149	1.261
	(0.009)	(0.061)	(0.009)	(0.061)	(0.009)	(0.061)	(0.009)	(0.061)
Loan age part 2	1.003	1.037	1.003	1.037	1.003	1.034	1.003	1.034
	(0.004)	(0.014)	(0.004)	(0.014)	(0.004)	(0.015)	(0.004)	(0.015)
Loan age part 3	0.998	0.990	0.997	0.990	0.998	0.990	0.997	0.990
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	1.006	0.997	1.005	0.997	1.006	0.997	1.005	0.997
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables								
The credit score part 1	3.848	0.000	--	0.000	3.810	0.000	--	0.000
	(2.599)	(0.000)		(0.000)	(2.535)	(0.000)		(0.000)
The credit score part 2	31.674	0.000	--	0.000	25.200	0.000	--	0.000
	(29.752)	(0.000)		(0.000)	(23.610)	(0.000)		(0.000)
The credit score part 3	2.673	0.005	--	0.005	2.575	0.030	--	0.029
	(3.392)	(0.021)		(0.021)	(3.364)	(0.123)		(0.121)
The credit score part 4	713.742	0.000	--	0.000	636.691	0.000	--	0.000
	(1,190.462)	(0.000)		(0.000)	(1,064.963)	(0.000)		(0.000)
The debt-to-income ratio variables								
DTI part 1	0.413	11.481	0.336	11.424	0.412	8.781	0.342	8.741
	(0.155)	(17.630)	(0.126)	(17.542)	(0.157)	(13.789)	(0.129)	(13.726)
DTI part 2	2.823	39.318	2.192	39.188	3.187	19.528	2.421	19.456
	(1.425)	(57.947)	(1.106)	(57.754)	(1.807)	(33.949)	(1.371)	(33.823)
DTI part 3	0.507	2.894	0.427	2.882	0.654	6.359	0.562	6.336
	(0.279)	(3.834)	(0.234)	(3.818)	(0.318)	(7.888)	(0.273)	(7.860)
DTI part 4	1.425	1.965	1.545	1.969	1.347	1.815	1.459	1.819
	(0.500)	(1.365)	(0.540)	(1.368)	(0.467)	(1.322)	(0.505)	(1.325)
The dummy for the number of units	0.543	0.916	0.567	0.916	0.542	0.742	0.564	0.742
	(0.126)	(0.463)	(0.132)	(0.463)	(0.126)	(0.432)	(0.131)	(0.432)
Time period variables								
Time period part 1	1.023	0.999	1.023	0.999	1.022	0.999	1.023	0.999
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	0.964	1.003	0.964	1.003	0.964	0.997	0.965	0.997
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	1.017	1.007	1.019	1.007	1.016	1.011	1.018	1.011
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.009)	(0.001)	(0.009)
	Number of mortgages 12,698				Number of mortgages 12,522			

In the model with the credit score used to explain both termination risks, the result of the indicator for a prepayment penalty supports H1. The odds for mortgages with a prepayment penalty are only 0.441 times as high as the same odds for mortgages without a prepayment penalty.

The effect of the value of the call option supports H2. When the value of the call option is below -0.025, the coefficient is 5.402 (the standard error is 0.893). When the value ranges between -0.025 and 0.035, the coefficient increases to 7.349 (the standard error is 0.911) and with ranges between 0.035 and 0.100, the coefficient again increases to 8.383 (the standard error is 0.683). However, when the value of the call option is above 0.100, the coefficient drops to 2.177 (the standard error is 0.424). The significantly positive effect supports the argument that when the value of the call option is “in the money,” households have more incentive to prepay their mortgages.

The monthly unemployment rate has a significantly positive effect on the 90-days-delinquency decision. When the unemployment rate is below 5.700, the odds of delinquency relative to continuity increase by 1.343 times with a 1 percent increase in the unemployment rate. When the unemployment rate is above 5.700, the odds of the delinquency relative to continuity increase by 1.194 times with a 1 percent increase in the unemployment rate. The effect of the unemployment rate on the 90-days-delinquency is very similar to its effect on the default.

The value of negative equity has a significantly positive relationship with the 90-days-delinquency risk. When negative equity is below 18.029, the odds of the 90-days-delinquency relative to continuity increase by 1.014 times with a 1 unit increase in negative equity. When ranges are between 18.029 and 37.334, the odds of the 90-days-delinquency relative to continuity increase by 1.023 times with a 1 unit increase in negative equity and in ranges between 37.334

and 62.503, the odds of the 90-days-delinquency relative to continuity increase by 1.018 times with a 1 unit increase in negative equity. However, when the negative equity is above 62.503, the odds of the 90-days-delinquency relative to continuity increase 1.003 times with a 1 unit increase in negative equity.

Another important variable to evaluate the effect of negative equity is the negative equity dummy. The result indicates the negative equity dummy is significantly and positively correlated with the 90-days-delinquency decision. The odds for mortgages with negative equity being delinquent instead of continued are 2.337 times as high as the same odds for mortgages with non-negative equity. Overall, the effect of negative equity on the 90-days-delinquency is very similar with that on the default. However, the effect of the negative equity dummy on the delinquency is slightly less than that on the default.

The relationship between the credit score and the 90-days-delinquency decision is strongly negative. Moreover, the credit score has a strongly positive effect on the prepayment decision. The coefficient is 1.347 for prepayment (the standard error is 0.675) and is -8.144 for the 90-days-delinquency (the standard error is 1.270) when the credit score is below 0.689. The coefficient increases to 3.455 for prepayment (the standard error is 0.939) and decreases to -9.572 for the 90-days-delinquency (the standard error is 2.291) when the credit score ranges between 0.689 and 0.737. The coefficient decreases to 0.983 for prepayment (the standard error is 1.269) and increases to -5.221 for the 90-days-delinquency (the standard error is 3.874) when the score ranges between 0.737 and 0.774. The coefficient dramatically increases to 6.571 for prepayment (the standard error is 1.668) and drops to -21.612 for the 90-days-delinquency (the standard error is 6.631) when the credit score is above 0.744. The results indicate both termination risks are sensitive to the high-level credit score. Therefore, households with high credit scores are more likely to prepay and less likely to be delinquent. Comparing the results of credit score in this model with those in the model using prepayment and default as dependent variables, the coefficients and the standard errors for the prepayment are very similar between the two models. However, the coefficients for the 90-days-delinquency are smaller than those for the default.

For the 90-days-delinquency risk, the debt-to-income ratio has a significantly positive effect. However, for the prepayment risk, the effect of the debt-to-income ratio is inconsistent. When the debt-to-income ratio ranges below 0.250 and between 0.340 and 0.420, its effect on the prepayment risk is negative. When the ratio ranges between 0.250 and 0.340 and above 0.420, its effect on the prepayment risk becomes positive. For the 90-days-delinquency risk, when the debt-to-income ratio is below 0.250, the coefficient is 2.441 and the standard error is 1.536. When the ratio ranges between 0.250 and 0.340, the coefficient increases to 3.672 and the standard error is 1.474. When the debt-to-income ratio ranges between 0.340 and 0.420, the coefficient drops to 0.675 and the standard error is 1.325, and when the ranges are above 0.420, the coefficient further decreases to 0.675 and the standard error is 0.695. The coefficients of the debt-to-income ratio for the 90-days-delinquency is slightly larger than those for the default.

The dummy for the number of units has a negative effect on both prepayment and the 90-days-delinquency, and the effect is significant for prepayment. The odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.543 times as high as the same odds for mortgages covering only one house unit. Moreover, The odds for mortgages covering more than one house unit being delinquent instead of continued are about 0.916 times as high as the same odds for mortgages covering only one house unit. The results show mortgages for larger houses are less likely to be prepaid or delinquent. The odds ratio for the prepayment is very similar between the models using different dependent variables. However, the odds ratio for the 90-days-delinquency is larger than that for the default.

Log loan size has a significantly positive effect on the prepayment; however, the effect is insignificant when loan size ranges between $149,941 and $210,029. On the other hand, its effect on 90-days-delinquency is insignificant. Furthermore, for size below $105,030, the odds of prepayment relative to continuity increase by 2.195 times with a 1 unit increase in log loan size. In this range, the odds of the 90-days-delinquency relative to continuity increase by 1.384 times with a 1 unit increase in log loan size. When size ranges between $105,030 and $149,941, the odds of prepayment relative to continuity increase by 2.308 times with a 1 unit increase in log loan size. In this range, the odds of the 90-days-delinquency relative to continuity decrease by 0.960 times with a 1 unit increase in log loan size. When size ranges between $149,941 and $210,029, the odds of prepayment relative to continuity increase by 1.245 times with a 1 unit increase in log loan size. In this range, the odds of the 90-days-delinquency relative to continuity decrease by 0.576 times with a 1 unit increase in log loan size. When log loan size is above $210,029, the odds of prepayment relative to continuity increase by 2.670 times with a 1 unit increase in log loan size. In this range, the odds of the 90-days-delinquency relative to continuity decrease by 0.795 times with a 1 unit increase in log loan size.

Loan age has a strongly positive relationship with the prepayment, but its effect is insignificantly negative when its value is between 23 and 44 months. When loan age is below 23 months, the effect of loan age on the 90-days-delinquency decision is significantly positive. However, when loan age is above 23 months, its effect on the 90-days-delinquency becomes significantly negative. When loan age is below 11 months, the odds of prepayment relative to continuity increase by 1.150 times with a 1 unit increase in loan age. In this range, the odds of the 90-days-delinquency instead of the continuity increase by 1.276 times with a 1 unit increase in loan age. When loan age ranges between 11 and 23 months, the odds of prepayment relative to continuity increase by 1.003 times with a 1 unit increase in loan age. In this range, the odds of the 90-days-delinquency relative to continuity increase by 1.037 times with a 1 unit increase in loan age. With age ranges between 23 and 44 months, the odds of prepayment relative to continuity rapidly decrease by 0.998 times with a 1 unit increase in loan age. In this range, the odds of the 90-days-delinquency relative to continuity decrease by 0.990 times with a 1 unit increase in loan age. However, when loan age is above 44 months, the odds of prepayment relative to continuity increase again by 1.006 times with a 1 unit increase in loan age and the odds ratio for the 90-days-delinquency remains the same as 0.997 in this range.

The time period in each model is also controlled. The results show mortgages are more likely to be prepaid from March 1999 to July 2003 and are less likely to be prepaid from July 2003 to November 2008. Prepayment increases again after November 2008. On the other hand, before January 2008, mortgages are less likely to be delinquent. However, starting from January 2010, mortgages become more likely to be delinquent.

The strongly negative effect of the credit score is still held in the model with the credit score explains only the 90-days-delinquency risk. Comparing the multinomial logit coefficients of the credit score for the 90-days-delinquency risk in this model with those in the previous model, results indicate the coefficients in this model are smaller than those in the previous model. However, the standard errors do not change. When the credit score is below 0.689, the coefficient is -8.144 and the standard error is 1.270. When the score ranges between 0.689 and 0.737, the coefficient decreases to -9.632 and the standard error is 2.491 and with ranges between 0.737 and 0.774, the coefficient increases to -5.238 and the standard error is 3.874.

However, when the credit score is above 0.744, the coefficient dramatically drops to -21.727 and the standard error is 6.631. These results indicate households with the higher credit scores are much less likely to have 90-days-delinquency.

Comparing the results of other explanatory variables in this model with those in the previous model, findings indicate the effect of the unemployment rate, negative equity, negative equity dummy, loan age, and time period on both termination risks remains the same. This also holds true for the effect of the log loan size and unit dummy on the 90-days-delinquency risk. The effect of the prepayment penalty still significantly negative, with the odds ratio slightly increase. The effect of the value of the call option is strongly positive, but the coefficients in this model are smaller than that in the previous model. The relationship between the debt-to-income ratio and the 90-days-delinquency risk is still positive. However, the coefficients in this model decrease about 0.004 with the same standard errors when the debt-to-income ratio is below 0.42 and increase about 0.003 with the same standard errors when the ratio is above 0.42. For the prepayment risk, the effect of the debt-to-income ratio is still inconsistent and the coefficients change considerably. The dummy for the number of units still has a significantly negative effect on the prepayment and the effect is stronger than that in the previous model. The effect of log loan size in most ranges is still significantly positive on the prepayment risk. Comparing the odds ratios in this model with those in the previous model, results indicate the odds ratios in this model are larger.

The second group of models includes negative equity as explanatory variables. Table 23 contains the coefficients and table 24 contains the odds ratios. The following explanations are based on the dataset with right censored mortgages and the results for the dataset with modified mortgages deleted can be found in tables 23 and 24. Comparing the results of this model with the results of the model using negative equity and the negative equity dummy, findings suggest the significance and the coefficients of all explanatory variables for the prepayment risk are the same. Therefore, the discussion below only focuses on the results for the 90-days-delinquency risk. The results of the model with the credit score for both termination risks is discussed first and then compared with the results of the model with the credit score only for the 90-days-delinquency risk.

[insert Tables 23 and 24 are here]

The third column in table 17 shows the monthly unemployment rate has a significantly positive effect on the 90-days-delinquency decision. When the unemployment rate is below 5.700, the odds of the 90-days-delinquency relative to continuity increase by 1.372 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of the 90-days-delinquency relative to continuity increase by 1.192 times with a 1 percent increase in the unemployment rate.

When the negative equity dummy is deleted from the model, the effect of negative equity is still significantly positive and most of the coefficients keep the same as those in the first model. However, when negative equity is below 37.334, the coefficients slightly change. The odds of the 90-days-delinquency relative to continuity increase by 1.065 times with a 1 unit increase in negative equity when negative equity is below 18.029, and the odds of the 90-days-delinquency relative to continuity increased by 1.014 times with a 1 unit increase in negative equity when its value ranges between 18.029 and 37.334.

The significantly negative effect of the credit score on the 90-days-delinquency risk still holds, but the coefficients are smaller in most ranges. The coefficient is -8.055 (the standard

Table 23. Coefficients for Phoenix: Prepayment and 90-days-delinquency are dependent variables, negative equity is one of explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency		The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	-0.819 (0.232)	--	-0.785 (0.232)	--	-0.823 (0.232)	--	-0.792 (0.232)	--
Call option variables								
Call option part 1	5.401 (0.893)	--	5.391 (0.892)	--	5.446 (0.893)	--	5.447 (0.892)	--
Call option part 2	7.349 (0.911)	--	7.214 (0.910)	--	7.424 (0.910)	--	7.291 (0.910)	--
Call option part 3	8.383 (0.683)	--	8.043 (0.682)	--	8.438 (0.684)	--	8.115 (0.683)	--
Call option part 4	2.177 (0.424)	--	1.667 (0.422)	--	2.285 (0.424)	--	1.798 (0.423)	--
The unemployment rate variables								
The unemployment rate part 1	--	0.316 (0.076)	--	0.316 (0.076)	--	0.309 (0.079)	--	0.309 (0.079)
The unemployment rate part 2	--	0.176 (0.052)	--	0.176 (0.052)	--	0.187 (0.054)	--	0.187 (0.054)
Negative equity variables								
Negative equity part 1	--	0.063 (0.007)	--	0.063 (0.007)	--	0.069 (0.008)	--	0.069 (0.008)
Negative equity part 2	--	0.014 (0.008)	--	0.014 (0.008)	--	0.015 (0.008)	--	0.015 (0.008)
Negative equity part 3	--	0.019 (0.005)	--	0.019 (0.005)	--	0.019 (0.005)	--	0.019 (0.005)
Negative equity part 4	--	0.003 (0.003)	--	0.003 (0.003)	--	0.004 (0.003)	--	0.004 (0.003)
Log loan size variables								
Log loan size part 1	0.786 (0.079)	0.369 (0.308)	0.704 (0.079)	0.368 (0.308)	0.787 (0.080)	0.413 (0.324)	0.707 (0.080)	0.413 (0.324)
Log loan size part 2	0.836 (0.125)	-0.024 (0.384)	0.809 (0.124)	-0.023 (0.384)	0.838 (0.121)	-0.279 (0.391)	0.811 (0.121)	-0.278 (0.391)
Log loan size part 3	0.219 (0.135)	-0.589 (0.348)	0.209 (0.135)	-0.589 (0.348)	0.218 (0.136)	-0.490 (0.371)	0.208 (0.136)	-0.490 (0.371)
Log loan size part 4	0.982 (0.094)	-0.233 (0.246)	0.990 (0.094)	-0.231 (0.246)	1.004 (0.094)	-0.282 (0.259)	1.012 (0.093)	-0.281 (0.259)
Loan age variables								
Loan age part 1	0.140 (0.008)	0.246 (0.048)	0.140 (0.008)	0.246 (0.048)	0.139 (0.008)	0.234 (0.048)	0.139 (0.008)	0.234 (0.048)
Loan age part 2	0.003 (0.004)	0.037 (0.014)	0.003 (0.004)	0.037 (0.014)	0.003 (0.004)	0.034 (0.014)	0.003 (0.004)	0.034 (0.014)
Loan age part 3	-0.002 (0.002)	-0.009 (0.005)	-0.003 (0.002)	-0.009 (0.005)	-0.002 (0.002)	-0.009 (0.005)	-0.003 (0.002)	-0.009 (0.005)
Loan age part 4	0.006 (0.001)	-0.003 (0.003)	0.005 (0.001)	-0.003 (0.003)	0.006 (0.001)	-0.003 (0.003)	0.005 (0.001)	-0.003 (0.003)
The credit score								

variables								
The credit score part 1	1.347	-8.055	--	-8.077	1.338	-8.182	--	-8.205
	(0.675)	(1.269)		(1.269)	(0.665)	(1.328)		(1.328)
The credit score part 2	3.455	-9.720	--	-9.780	3.227	-9.610	--	-9.668
	(0.939)	(2.290)		(2.290)	(0.937)	(2.408)		(2.408)
The credit score part 3	0.983	-5.303	--	-5.320	0.946	-3.579	--	-3.595
	(1.269)	(3.871)		(3.871)	(1.306)	(4.131)		(4.131)
The credit score part 4	6.570	-21.989	--	-22.104	6.456	-21.765	--	-21.880
	(1.668)	(6.625)		(6.625)	(1.673)	(6.773)		(6.773)
The debt-to-income ratio variables								
DTI part 1	-0.884	2.457	-1.091	2.452	-0.886	2.187	-1.074	2.182
	(0.376)	(1.535)	(0.374)	(1.535)	(0.380)	(1.570)	(0.378)	(1.570)
DTI part 2	1.038	3.665	0.785	3.662	1.159	2.967	0.884	2.964
	(0.505)	(1.472)	(0.505)	(1.472)	(0.567)	(1.737)	(0.566)	(1.737)
DTI part 3	-0.678	1.131	-0.851	1.127	-0.424	1.905	-0.576	1.902
	(0.550)	(1.324)	(0.549)	(1.324)	(0.486)	(1.240)	(0.485)	(1.240)
DTI part 4	0.354	0.742	0.435	0.744	0.298	0.666	0.378	0.668
	(0.351)	(0.694)	(0.350)	(0.694)	(0.347)	(0.728)	(0.346)	(0.728)
The dummy for the number of units	-0.610	-0.097	-0.567	-0.097	-0.613	-0.305	-0.573	-0.306
	(0.233)	(0.505)	(0.232)	(0.505)	(0.233)	(0.582)	(0.232)	(0.582)
Time period variables								
Time period part 1	0.022	0.000	0.023	0.000	0.022	0.000	0.023	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	-0.037	0.003	-0.036	0.003	-0.036	-0.003	-0.035	-0.003
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	0.016	0.006	0.019	0.006	0.015	0.010	0.018	0.010
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)

Number of mortgages 12,698 Number of mortgages 12,522

Table 24. Odds ratios for Phoenix: Prepayment and 90-days-delinquency are dependent variables, negative equity is one of explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency		The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	0.441	--	0.456	--	0.439	--	0.453	--
	(0.102)		(0.106)		(0.102)		(0.105)	
Call option variables								
Call option part 1	221.732	--	219.441	--	231.798	--	232.176	--
	(197.927)		(195.733)		(206.880)		(207.060)	
Call option part 2	1,555.223	--	1,358.698	--	1,675.090	--	1,466.657	--
	(1,416.388)		(1,236.806)		(1,525.127)		(1,334.691)	
Call option part 3	4,372.500	--	3,111.618	--	4,617.703	--	3,345.660	--
	(2,985.200)		(2,120.909)		(3,158.495)		(2,284.790)	
Call option part 4	8.819	--	5.297	--	9.824	--	6.036	--
	(3.739)		(2.236)		(4.170)		(2.551)	
The unemployment rate variables								
The unemployment rate part 1	--	1.372	--	1.372	--	1.363	--	1.363
		(0.104)		(0.104)		(0.107)		(0.108)

The unemployment rate part 2	--	1.192	--	1.192	--	1.205	--	1.205
		(0.061)		(0.061)		(0.065)		(0.065)
Negative equity variables								
Negative equity part 1	--	1.065	--	1.065	--	1.071	--	1.071
		(0.008)		(0.008)		(0.008)		(0.008)
Negative equity part 2	--	1.014	--	1.014	--	1.015	--	1.015
		(0.008)		(0.008)		(0.009)		(0.009)
Negative equity part 3	--	1.019	--	1.019	--	1.020	--	1.020
		(0.005)		(0.005)		(0.006)		(0.006)
Negative equity part 4	--	1.003	--	1.003	--	1.004	--	1.004
		(0.003)		(0.003)		(0.003)		(0.003)
Log loan size variables								
Log loan size part 1	2.195	1.446	2.022	1.445	2.196	1.511	2.028	1.511
	(0.174)	(0.446)	(0.159)	(0.446)	(0.176)	(0.489)	(0.161)	(0.489)
Log loan size part 2	2.308	0.976	2.245	0.977	2.311	0.756	2.250	0.757
	(0.287)	(0.375)	(0.279)	(0.375)	(0.281)	(0.296)	(0.273)	(0.296)
Log loan size part 3	1.245	0.555	1.232	0.555	1.244	0.613	1.231	0.613
	(0.168)	(0.193)	(0.166)	(0.193)	(0.169)	(0.227)	(0.168)	(0.227)
Log loan size part 4	2.670	0.792	2.691	0.794	2.730	0.754	2.752	0.755
	(0.252)	(0.195)	(0.253)	(0.195)	(0.255)	(0.195)	(0.257)	(0.195)
Loan age variables								
Loan age part 1	1.150	1.279	1.150	1.279	1.149	1.264	1.149	1.264
	(0.009)	(0.061)	(0.009)	(0.061)	(0.009)	(0.061)	(0.009)	(0.061)
Loan age part 2	1.003	1.038	1.003	1.038	1.003	1.035	1.003	1.035
	(0.004)	(0.014)	(0.004)	(0.014)	(0.004)	(0.015)	(0.004)	(0.015)
Loan age part 3	0.998	0.991	0.997	0.991	0.998	0.991	0.997	0.991
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	1.006	0.997	1.005	0.997	1.006	0.997	1.005	0.997
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables								
The credit score part 1	3.847	0.000	--	0.000	3.810	0.000	--	0.000
	(2.599)	(0.000)		(0.000)	(2.535)	(0.000)		(0.000)
The credit score part 2	31.666	0.000	--	0.000	25.194	0.000	--	0.000
	(29.745)	(0.000)		(0.000)	(23.605)	(0.000)		(0.000)
The credit score part 3	2.673	0.005	--	0.005	2.575	0.028	--	0.027
	(3.392)	(0.019)		(0.019)	(3.364)	(0.115)		(0.113)
The credit score part 4	713.723	0.000	--	0.000	636.666	0.000	--	0.000
	(1,190.432)	(0.000)		(0.000)	(1,064.923)	(0.000)		(0.000)
The debt-to-income ratio variables								
DTI part 1	0.413	11.670	0.336	11.611	0.412	8.904	0.342	8.863
	(0.155)	(17.916)	(0.126)	(17.826)	(0.157)	(13.980)	(0.129)	(13.915)
DTI part 2	2.823	39.066	2.192	38.936	3.187	19.439	2.421	19.367
	(1.425)	(57.517)	(1.106)	(57.326)	(1.807)	(33.757)	(1.372)	(33.632)
DTI part 3	0.507	3.099	0.427	3.086	0.654	6.720	0.562	6.696
	(0.279)	(4.104)	(0.235)	(4.088)	(0.318)	(8.335)	(0.273)	(8.305)
DTI part 4	1.425	2.100	1.545	2.104	1.347	1.947	1.459	1.951
	(0.500)	(1.457)	(0.540)	(1.460)	(0.467)	(1.417)	(0.505)	(1.420)
The dummy for the number of units	0.543	0.908	0.567	0.908	0.542	0.737	0.564	0.737
	(0.126)	(0.459)	(0.132)	(0.459)	(0.126)	(0.429)	(0.131)	(0.429)
Time period variables								
Time period part 1	1.023	1.000	1.023	1.000	1.022	1.000	1.023	1.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	0.964	1.003	0.964	1.003	0.964	0.997	0.965	0.997
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	1.017	1.006	1.019	1.006	1.016	1.010	1.018	1.010
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.009)	(0.001)	(0.009)

Number of mortgages 12,698

Number of mortgages 12,522

error is 1.269) when the credit score is below 0.689 and the coefficient decreases to -9.720 (the standard error is 2.290) when the credit score ranges between 0.689 and 0.737. The coefficient further increases to -5.303 (the standard error is 3.871) when the credit score ranges between 0.737 and 0.774 and the coefficient dramatically drops to -21.989 (the standard error is 6.625) when the credit score is above 0.744. The results indicate households with the higher credit scores are much less likely to have 90-days-delinquency.

The effect of the debt-to-income ratio on the 90-days-delinquency risk is positive, but it is only significant when the debt-to-income ratio ranges between 0.250 and 0.340. When the ratio is below 0.250, the coefficient is 2.457 (the standard error is 1.535), and when the ratio ranges between 0.250 and 0.340, the coefficient increases to 3.665 (the standard error is 1.472). When the debt-to-income ratio ranges between 0.340 and 0.420, the coefficient drops to 1.131 (the standard error is 1.324), and when the ratio is above 0.420, the coefficient further decreases to 0.742 (the standard error is 0.694).

The dummy for the number of units has a negative effect on the 90-days-delinquency risk. The odds for mortgages covering more than one house unit being delinquent instead of continued are about 0.908 times as high as the same odds for mortgages covering only one house unit. The results show mortgages for larger houses are less likely to be delinquent than mortgages for smaller houses.

Log loan size has a negative effect on 90-days-delinquency. However, when loan size is below $105,030, its effect becomes insignificantly positive. When loan size is below $105,030, the odds of the 90-days-delinquency relative to continuity increase by 1.446 times with a 1 unit increase in log loan size. When size ranges between $105,030 and $149,941, the odds of the 90-days-delinquency relative to continuity decrease by 0.976 times with a 1 unit increase in log loan size, and when its value ranges between $149,941 and $210,029, the odds of the 90-days-delinquency relative to continuity decrease by 0.555 times with a 1 unit increase in log loan size when the size is above $210,029, the odds of the 90-days-delinquency relative to continuity decrease by 0.792 times with a 1 unit increase in log loan size.

Loan age has a significantly positive effect on the 90-days-delinquency risk when the age is below 23 months, and its effect becomes significantly negative when the age is above 23 months. When loan age is below 11, the odds of the 90-days-delinquency relative to continuity increase by 1.279 times with a 1 unit increase in loan age. When age ranges between 11 and 23 months, the odds of the 90-days-delinquency relative to continuity increase by 1.038 times with a 1 unit increase in loan age, and when it ranges between 23 and 44 months, the odds of the 90-days-delinquency relative to continuity decrease by 0.991 times with a 1 unit increase in loan age. When loan age is above 44 months, the odds of the 90-days-delinquency relative to continuity decrease by 0.997 times with a 1 unit increase in loan age.

The strongly negative effect of the credit score still holds in the model with the credit score explains only the 90-days-delinquency risk. Comparing the coefficients of the credit score for the default risk in this model with those in the model with the credit score used for both termination risks, findings indicate the coefficients in this model are smaller than those in the previous model. However, the standard errors do not change. When the credit score is below 0.689, the coefficient is -8.077 (the standard error is 1.269), and when the score ranges between 0.689 and 0.737, the coefficient decreases to -9.780 (the standard error is 2.290). When the credit score ranges between 0.737 and 0.774, the coefficient increases to -5.320 (the standard error is 3.871), and when the score is above 0.744, the coefficient dramatically drops to -22.104 (the standard error is 6.625). This result indicates households with high credit scores are much less likely to default.

The results of the other explanatory variables for the 90-days-delinquency risk in this model are compared with those in the model with the credit score used for both termination risks, and findings indicate the effect of the unemployment rate, negative equity, negative equity dummy, log loan size, loan age, unit dummy and time period remains the same. However, the coefficients of the debt-to-income ratio slightly decrease. When the debt-to-income ratio is below 0.42, the coefficients in this model decrease about 0.004 with the same standard errors, and when the ratio is above 0.42 the coefficient increases about 0.002 with the same standard errors.

The third group of models includes negative equity, negative equity dummy, and original loan-to-value as explanatory variables. Table 25 contains the coefficients and table 26 contains the odds ratios. The following explanations are based on the dataset with right censored mortgages and the results for the dataset with modified mortgages deleted can be found in tables 25 and 26. The results of this model are compared with the results of the model using negative equity and negative equity dummy as explanatory variables, and findings indicate the significance and coefficients of all explanatory variables for the prepayment risk are the same. Therefore, only the results for the 90-days-delinquency risk are discussed. The results of the model with the credit score for both termination risks are reviewed and then compared with the results of the model with the credit score only for the 90-days-delinquency risk.

[insert Tables 25 and 26 are here]

Table 25. Coefficients for Phoenix: Prepayment and 90-days-delinquency are dependent variables, negative equity, the negative equity dummy and original loan-to-value are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency		The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	-0.819 (0.232)	--	-0.785 (0.232)	--	-0.823 (0.232)	--	-0.792 (0.232)	--
Call option variables								
Call option part 1	5.401 (0.893)	--	5.391 (0.892)	--	5.446 (0.892)	--	5.447 (0.892)	--
Call option part 2	7.350 (0.911)	--	7.215 (0.910)	--	7.424 (0.910)	--	7.292 (0.910)	--
Call option part 3	8.382 (0.683)	--	8.042 (0.682)	--	8.437 (0.684)	--	8.115 (0.683)	--
Call option part 4	2.179 (0.424)	--	1.670 (0.422)	--	2.287 (0.425)	--	1.800 (0.423)	--
The unemployment rate variables								
The unemployment rate part 1	--	0.308 (0.076)	--	0.308 (0.076)	--	0.302 (0.079)	--	0.302 (0.079)
The unemployment rate part 2	--	0.189 (0.051)	--	0.189 (0.051)	--	0.200 (0.054)	--	0.200 (0.054)
Negative equity variables								
Negative equity part 1	--	0.013 (0.014)	--	0.013 (0.014)	--	0.018 (0.014)	--	0.018 (0.014)
Negative equity part 2	--	0.020 (0.008)	--	0.020 (0.008)	--	0.020 (0.009)	--	0.020 (0.009)
Negative equity part 3	--	0.018 (0.005)	--	0.018 (0.005)	--	0.018 (0.006)	--	0.018 (0.006)
Negative equity part 4	--	-0.000 (0.003)	--	-0.000 (0.003)	--	0.001 (0.003)	--	0.001 (0.003)
The negative equity dummy	--	0.780 (0.197)	--	0.780 (0.197)	--	0.782 (0.211)	--	0.782 (0.211)
Original loan-to-value variables								
Original LTV part 1	--	2.295 (0.910)	--	2.295 (0.910)	--	2.644 (0.998)	--	2.644 (0.998)
Original LTV part 2	--	0.560 (1.103)	--	0.561 (1.103)	--	0.622 (1.174)	--	0.623 (1.174)
Original LTV part 3	--	-7.434 (4.980)	--	-7.433 (4.980)	--	-5.996 (5.240)	--	-5.995 (5.240)
Original LTV part 4	--	3.362 (0.699)	--	3.362 (0.699)	--	3.446 (0.725)	--	3.447 (0.725)
Original LTV part 5	--	2.522 (4.248)	--	2.517 (4.248)	--	1.460 (4.554)	--	1.453 (4.554)
Log loan size variables								
Log loan size part 1	0.786 (0.079)	0.134 (0.313)	0.704 (0.079)	0.134 (0.313)	0.787 (0.080)	0.149 (0.329)	0.707 (0.080)	0.149 (0.329)
Log loan size part 2	0.836	0.009	0.809	0.010	0.838	-0.227	0.811	-0.226

	(0.125)	(0.384)	(0.124)	(0.384)	(0.121)	(0.392)	(0.121)	(0.392)
Log loan size part 3	0.219	-0.536	0.209	-0.536	0.218	-0.446	0.208	-0.446
	(0.135)	(0.347)	(0.135)	(0.347)	(0.136)	(0.369)	(0.136)	(0.369)
Log loan size part 4	0.982	-0.057	0.990	-0.055	1.004	-0.090	1.012	-0.089
	(0.094)	(0.245)	(0.094)	(0.245)	(0.094)	(0.258)	(0.093)	(0.258)
Loan age variables								
Loan age part 1	0.140	0.240	0.140	0.240	0.139	0.228	0.139	0.228
	(0.008)	(0.048)	(0.008)	(0.048)	(0.008)	(0.048)	(0.008)	(0.048)
Loan age part 2	0.003	0.040	0.003	0.040	0.003	0.037	0.003	0.037
	(0.004)	(0.014)	(0.004)	(0.014)	(0.004)	(0.014)	(0.004)	(0.014)
Loan age part 3	-0.002	-0.007	-0.003	-0.007	-0.002	-0.006	-0.003	-0.006
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	0.006	-0.005	0.005	-0.005	0.006	-0.005	0.005	-0.005
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables								
The credit score part 1	1.347	-8.436	--	-8.458	1.337	-8.552	--	-8.576
	(0.675)	(1.281)		(1.281)	(0.665)	(1.339)		(1.339)
The credit score part 2	3.455	-8.847	--	-8.907	3.227	-8.743	--	-8.801
	(0.939)	(2.299)		(2.299)	(0.937)	(2.417)		(2.417)
The credit score part 3	0.983	-5.198	--	-5.214	0.946	-3.544	--	-3.561
	(1.269)	(3.878)		(3.878)	(1.306)	(4.138)		(4.138)
The credit score part 4	6.571	-21.038	--	-21.155	6.457	-20.798	--	-20.915
	(1.668)	(6.626)		(6.626)	(1.673)	(6.776)		(6.776)
The debt-to-income ratio variables								
DTI part 1	-0.884	2.364	-1.091	2.359	-0.886	2.048	-1.074	2.044
	(0.376)	(1.534)	(0.374)	(1.534)	(0.380)	(1.568)	(0.378)	(1.568)
DTI part 2	1.038	3.426	0.785	3.423	1.159	2.770	0.884	2.766
	(0.505)	(1.471)	(0.505)	(1.471)	(0.567)	(1.737)	(0.566)	(1.737)
DTI part 3	-0.678	1.129	-0.851	1.125	-0.424	1.876	-0.576	1.873
	(0.550)	(1.325)	(0.549)	(1.325)	(0.486)	(1.243)	(0.485)	(1.243)
DTI part 4	0.354	0.546	0.435	0.548	0.297	0.429	0.377	0.431
	(0.351)	(0.698)	(0.350)	(0.698)	(0.347)	(0.733)	(0.346)	(0.733)
The dummy for the number of units	-0.611	-0.028	-0.567	-0.028	-0.613	-0.228	-0.573	-0.228
	(0.233)	(0.506)	(0.232)	(0.506)	(0.233)	(0.583)	(0.232)	(0.583)
Time period variables								
Time period part 1	0.022	0.002	0.023	0.002	0.022	0.002	0.023	0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	-0.037	0.002	-0.036	0.002	-0.036	-0.004	-0.035	-0.004
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	0.016	0.007	0.019	0.007	0.015	0.011	0.018	0.011
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)
	Number of mortgages 12,698				Number of mortgages 12,522			

Table 26. Odds ratios for Phoenix: Prepayment and 90-days-delinquency are dependent variables, negative equity, the negative equity dummy and original loan-to-value are explanatory variables

	#1 dataset with right censored mortgages				#2 dataset with modified mortgages deleted			
	The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency		The credit score for both the prepayment and the 90-days-delinquency		The credit score only for 90-days-delinquency	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency

Prepayment penalty	0.441	--	0.456	--	0.439	--	0.453	--
	(0.102)		(0.106)		(0.102)		(0.105)	
Call option variables								
Call option part 1	221.725	--	219.437	--	231.779	--	232.157	--
	(197.920)		(195.729)		(206.862)		(207.043)	
Call option part 2	1,556.495	--	1,359.846	--	1,676.561	--	1,467.991	--
	(1,417.547)		(1,237.852)		(1,526.468)		(1,335.906)	
Call option part 3	4,369.498	--	3,109.269	--	4,614.286	--	3,342.926	--
	(2,983.157)		(2,119.312)		(3,156.165)		(2,282.928)	
Call option part 4	8.840	--	5.310	--	9.847	--	6.050	--
	(3.748)		(2.242)		(4.180)		(2.557)	
The unemployment rate variables								
The unemployment rate part 1	--	1.361	--	1.361	--	1.353	--	1.353
		(0.103)		(0.103)		(0.107)		(0.107)
The unemployment rate part 2	--	1.207	--	1.207	--	1.221	--	1.221
		(0.062)		(0.062)		(0.066)		(0.066)
Negative equity variables								
Negative equity part 1	--	1.013	--	1.013	--	1.018	--	1.018
		(0.014)		(0.014)		(0.015)		(0.015)
Negative equity part 2	--	1.020	--	1.020	--	1.020	--	1.020
		(0.009)		(0.009)		(0.009)		(0.009)
Negative equity part 3	--	1.018	--	1.018	--	1.018	--	1.018
		(0.005)		(0.005)		(0.006)		(0.006)
Negative equity part 4	--	1.000	--	1.000	--	1.001	--	1.001
		(0.003)		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	2.182	--	2.181	--	2.186	--	2.185
		(0.431)		(0.430)		(0.461)		(0.461)
Original loan-to-value variables								
Original LTV part 1	--	9.925	--	9.922	--	14.074	--	14.070
		(9.035)		(9.033)		(14.050)		(14.046)
Original LTV part 2	--	1.751	--	1.752	--	1.863	--	1.864
		(1.932)		(1.933)		(2.186)		(2.187)
Original LTV part 3	--	0.001	--	0.001	--	0.002	--	0.002
		(0.003)		(0.003)		(0.013)		(0.013)
Original LTV part 4	--	28.841	--	28.861	--	31.376	--	31.395
		(20.174)		(20.187)		(22.733)		(22.746)
Original LTV part 5	--	12.451	--	12.390	--	4.305	--	4.278
		(52.893)		(52.633)		(19.605)		(19.480)
Log loan size variables								
Log loan size part 1	2.195	1.143	2.023	1.143	2.196	1.161	2.028	1.161
	(0.174)	(0.358)	(0.159)	(0.357)	(0.176)	(0.381)	(0.161)	(0.381)
Log loan size part 2	2.308	1.009	2.245	1.010	2.311	0.797	2.250	0.798
	(0.287)	(0.388)	(0.279)	(0.388)	(0.281)	(0.312)	(0.273)	(0.312)
Log loan size part 3	1.245	0.585	1.232	0.585	1.244	0.640	1.231	0.640
	(0.168)	(0.203)	(0.166)	(0.203)	(0.169)	(0.237)	(0.168)	(0.237)
Log loan size part 4	2.670	0.945	2.691	0.946	2.730	0.914	2.751	0.915
	(0.252)	(0.232)	(0.253)	(0.232)	(0.255)	(0.236)	(0.257)	(0.236)
Loan age variables								
Loan age part 1	1.150	1.271	1.150	1.272	1.149	1.256	1.149	1.256
	(0.009)	(0.061)	(0.009)	(0.061)	(0.009)	(0.061)	(0.009)	(0.061)
Loan age part 2	1.003	1.040	1.003	1.040	1.003	1.037	1.003	1.037
	(0.004)	(0.014)	(0.004)	(0.014)	(0.004)	(0.015)	(0.004)	(0.015)
Loan age part 3	0.998	0.993	0.997	0.993	0.998	0.994	0.997	0.994
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	1.006	0.995	1.005	0.995	1.006	0.995	1.005	0.995
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables								
The credit score part 1	3.846	0.000	--	0.000	3.809	0.000	--	0.000
	(2.598)	(0.000)		(0.000)	(2.534)	(0.000)		(0.000)

The credit score part 2	31.669	0.000	--	0.000	25.198	0.000	--	0.000
	(29.747)	(0.000)		(0.000)	(23.608)	(0.000)		(0.000)
The credit score part 3	2.673	0.006	--	0.005	2.575	0.029	--	0.028
	(3.391)	(0.021)		(0.021)	(3.363)	(0.120)		(0.118)
The credit score part 4	714.151	0.000	--	0.000	637.075	0.000	--	0.000
	(1,191.148)	(0.000)		(0.000)	(1,065.610)	(0.000)		(0.000)
The debt-to-income ratio variables								
DTI part 1	0.413	10.633	0.336	10.581	0.412	7.756	0.342	7.721
	(0.155)	(16.314)	(0.126)	(16.234)	(0.157)	(12.159)	(0.129)	(12.105)
DTI part 2	2.823	30.749	2.192	30.649	3.187	15.960	2.421	15.901
	(1.425)	(45.240)	(1.106)	(45.094)	(1.807)	(27.725)	(1.372)	(27.623)
DTI part 3	0.507	3.093	0.427	3.081	0.655	6.528	0.562	6.505
	(0.279)	(4.098)	(0.235)	(4.082)	(0.318)	(8.112)	(0.273)	(8.084)
DTI part 4	1.424	1.727	1.545	1.731	1.346	1.536	1.458	1.539
	(0.499)	(1.206)	(0.540)	(1.209)	(0.467)	(1.126)	(0.505)	(1.129)
The dummy for the number of units	0.543	0.972	0.567	0.972	0.542	0.796	0.564	0.796
	(0.126)	(0.491)	(0.132)	(0.491)	(0.126)	(0.464)	(0.131)	(0.464)
Time period variables								
Time period part 1	1.023	1.002	1.023	1.002	1.022	1.002	1.023	1.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	0.964	1.002	0.964	1.002	0.964	0.996	0.965	0.996
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.010)	(0.001)	(0.010)
Time period part 3	1.017	1.007	1.019	1.007	1.016	1.011	1.018	1.011
	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)	(0.009)	(0.001)	(0.009)
	Number of mortgages 12,698				Number of mortgages 12,522			

The monthly unemployment rate has a significantly positive effect on the 90-days-delinquency decision. When the unemployment rate is below 5.700, the odds of the 90-days-delinquency relative to continuity increase by 1.361 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of the 90-days-delinquency relative to continuity increase by 1.207 times with a 1 percent increase in the unemployment rate.

When original loan-to-value is added into the model, the effect of negative equity is still significantly positive in most ranges. However, when negative equity is above 62.503, the effect becomes insignificantly negative. The odds of the 90-days-delinquency relative to continuity increase by 1.013 times with a 1 unit increase in negative equity when negative equity is below 18.029, and the odds of the 90-days-delinquency relative to continuity increase by 1.020 times with a 1 unit increase in negative equity when its value ranges between 18.029 and 37.334. The odds of the 90-days-delinquency relative to continuity increase by 1.018 times with a 1 unit increase in negative equity when negative equity ranges between 37.334 and 62.503, and the odds of the 90-days-delinquency relative to continuity will not change with a 1 unit increase in negative equity when its value is above 62.503.

The negative equity dummy is significantly and positively correlated with the 90-days-delinquency decision. The odds for mortgages with negative equity being delinquent instead of continued are 2.182 times as high as the same odds for mortgages with non-negative equity.

In most ranges, the effect of original loan-to-value is significantly positive. However, when its value ranges between 0.780 and 0.800, the effect becomes insignificantly negative. When original loan-to-value is below 0.650, the coefficient is 2.295 (the standard error is 0.901), and when the value ranges between 0.650 and 0.780, the coefficient decreases to 0.560 (the standard error is 1.103). In ranges between 0.780 and 0.800, the coefficient becomes insignificantly negative, which is -7.434 (the standard error is 4.980), and when ranges are between 0.800 and 0.950, the coefficient becomes significantly positive again, which is 3.362 (the standard error is 0.699). When original loan-to-value is above 0.950, the coefficient decreases slightly to 2.522 (the standard error is 4.248).

The relationship between the credit score and the 90-days-delinquency is significantly negative. The coefficient is -8.436 (the standard error is 1.281) when the credit score is below 0.689 and the coefficient decreases to -8.847 (the standard error is 2.299) when the score ranges between 0.689 and 0.737. The coefficient increases to -5.198 (the standard error is 3.878) when it ranges between 0.737 and 0.774 and the coefficient dramatically drops to -21.038 (the standard error is 6.626) when the credit score is above 0.744. The results indicate households with high credit scores are much less likely to have 90-days-delinquency.

The effect of the debt-to-income ratio on the 90-days-delinquency risk is positive but it is only significant when the debt-to-income ratio ranges between 0.250 and 0.340. The dummy for the number of units has a negative effect on the 90-days-delinquency risk. The odds for mortgages covering more than one house unit being delinquent instead of continued are about 0.972 times as high as the same odds for mortgages covering only one house unit. The results show mortgages for larger houses are less likely to be delinquent than mortgages for smaller houses.

The effect of log loan size on 90-days-delinquency risk become insignificantly positive when its value is below \$149,941, and the effect becomes insignificantly negative when its value is above \$149,941.

Loan age has a significantly positive effect on the 90-days-delinquency risk when the age is below 23 months, and its effect becomes significantly negative when the age is above 23 months. When loan age is below 11 months, the odds of the 90-days-delinquency relative to continuity increase by 1.271 times with a 1 unit increase in loan age. When age ranges between 11 and 23 months, the odds of the 90-days-delinquency relative to continuity increase by 1.040 times with a 1 unit increase in loan age. When loan age ranges between 23 and 44 months, the odds of the 90-days-delinquency relative to continuity decrease by 0.993 times with a 1 unit increase in loan age, and when age is above 44 months, the odds of the 90-days-delinquency relative to continuity decrease by 0.995 times with a 1 unit increase in loan age.

The strongly negative effect of the credit score still holds in the model with the credit score explains only the 90-days-delinquency risk. Comparing the coefficients of the credit score for the default risk in this model with those in the model with the credit score used for both termination risks, findings indicate the coefficients in this model are smaller than those in the previous model. However, the standard errors do not change. When the credit score is below 0.689, the coefficient is -8.907 (the standard error is 2.299), and when the score ranges between 0.689 and 0.737, the coefficient decreases to -8.907 (the standard error is 2.299). When the credit score ranges between 0.737 and 0.774, the coefficient increases to -5.214 (the standard error is 3.878), and when it is above 0.744, the coefficient dramatically drops to -21.155 (the standard error is 6.626). This result indicates households with high credit scores are much less likely to default.

The results of other explanatory variables for the 90-days-delinquency risk in this model are compared with those in the model with the credit score used for both termination risks, and findings show the effect of the unemployment rate, negative equity dummy, original loan-to-value, log loan size, loan age, unit dummy and time period remains the same. However, the coefficients of the debt-to-income ratio slightly decrease. When the debt-to-income ratio is below 0.42, the coefficients in this model decrease about 0.004 with the same standard errors, and when the ratio is above 0.42 the coefficient increases about 0.002 with the same standard errors.

Table 27 contains the Pseudo R-square and the BIC for these twenty-four models. The evaluation is separated into two groups based on the two combinations of the dependent variables. This discussion only focuses on the dataset with right censored mortgages and the evaluation for the dataset with modified mortgages deleted can be found in table 27. In the case where prepayment and default are used as dependent variables, the upper part of column 2 in table 27 indicates that when the credit score is used in both termination risks, the Pseudo R-squares are the same among models with the different specifications. This means the predictive power is the same among models. The upper part of column 4 in table 27 shows when the credit score is used for explaining only the default risk, the Pseudo R-squares are the same among models with the different specifications. However, comparing the BIC among models with different specifications, findings indicate the models containing negative equity, the negative equity dummy and original loan-to-value as explanatory variables have the largest BIC. This means that this model fits the data better. In the case where prepayment and 90-days-delinquency are used as dependent variables, the lower part of column 2 in table 27 indicates when the credit score is used in both termination risks, the Pseudo R-squares are the same among models with the different specifications. This means the predictive power is the same among models. The lower part of column 4 in table 27 shows when the credit score is used for explaining only the 90-days-delinquency risk, the Pseudo R-squares are the same among models with the different specifications. However, comparing the BIC among models with the different specifications, results indicate the models containing negative equity and negative equity dummy as explanatory variables have the largest BIC. This means this model fits the data better.

[insert Table 27 is here]

Since the samples size in Phoenix is the largest and the effect of the explanatory variables supports most of the hypothesis made in the previous literature, this paper only explains the results for Phoenix in detail. The results for Miami, Tampa, Detroit and Las Vegas are briefly explained and the detail information of results can be found in table 28 through table 47. The important results for these four MSAs are compared with the results for Phoenix and the comparison is presented in the following section.

The Comparison Results for Miami, Tampa, Detroit and Las Vegas

Table 28 through table 47 presents the estimated results for Miami, Tampa, Detroit and Las Vegas. For these MSAs, six results are listed in detail including two sets of dependent variables: prepayment and default, and prepayment and 90-days-delinquency. For these two sets of dependent variables, the best three specifications[10] are estimated based on the dataset with

10 For Miami, when the prepayment and the default are used as dependent variables, the best specification includes the negative equity dummy and original loan-to-value. The second best includes negative equity and original loan-to-value, and the third best includes the negative equity dummy. In the case when the prepayment and the 90-days-delinquency are used as dependent variables, the best specification includes the negative equity dummy and original loan-to-value. The second best includes the negative equity dummy and the third best includes negative equity and original loan-to-value.

In Tampa, for both sets of dependent variables, the best specification includes the negative equity dummy and the second best includes the negative equity dummy and original loan-to-value. The third best specification is different for models using a different set of dependent variables. When the prepayment and the default are used, the third best specification includes negative equity and the negative equity dummy. When the prepayment and the 90-days-delinquency are used, the third best specification includes the value of negative equity.

In Detroit, for two sets of dependent variables, the best specification includes original loan-to-value, the second best includes the negative equity dummy and original loan-to-value, and the third best includes negative equity and original loan-to-value.

In Las Vegas, the best specification includes negative equity, the second best includes negative equity and the negative equity dummy, and the third best includes negative equity and original loan-to-value. These specifications are used for two sets of dependent variables.

Table 27. Results comparison in Phoenix

Prepayment and Default are Dependent Variables

		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and default		The credit score for default		The credit score for prepayment and default		The credit score for default	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	negative equity, original LTV and the negative equity dummy	0.9154	-45021.5	0.9153	-45078	0.9144	-44899.4	0.9142	-44949.9
2	negative equity and the negative equity dummy	0.9154	-45021.9	0.9153	-45078.5	0.9143	-44900.4	0.9142	-44950.9
3	original LTV and the negative equity dummy	0.9154	-45023.7	0.9153	-45080.2	0.9143	-44903.1	0.9142	-44953.6

Prepayment and 90-days-delinquency are Dependent Variables

		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and 90-days-delinquency		The credit score for 90-days-delinquency		The credit score for prepayment and 90-days-delinquency		The credit score for 90-days-delinquency	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	negative equity and the negative equity dummy	0.9126	-45741.6	0.9124	-45789.4	0.9126	-45073.9	0.9124	-45116.6
2	negative equity	0.9125	-45745	0.9124	-45792.8	0.9126	-45076.4	0.9124	-45119.1
3	negative equity, original LTV and the negative equity dummy	0.9126	-45745.6	0.9125	-45793.4	0.9126	-45077.2	0.9125	-45119.9

mortgages being right censored one month before the modification date. Moreover, after comparing the BIC, the models with the credit score only for the default/90-days-delinquency risk are presented for Miami and Las Vegas and the models with the credit score for both termination risks are presented for Tampa and Detroit.

[insert Tables 28 through 47 here]

The important results for each MSAs show that for Miami, Tampa, and Las Vegas, the negative coefficient of the indicator for a prepayment penalty support H1 that the mortgages with a prepayment penalty are less likely to be prepaid compared with those without any penalty. However, the coefficient is insignificant for Detroit. This means that the prepayment penalty is not a determinant factor for the prepayment risk in Detroit.

For all MSAs, the significantly positive coefficients of the value of the call option support H2 that when the value of the call option is “in the money,” households have more incentive to prepay their mortgages.

For Miami and Tampa, the positive effect of the unemployment rate supports H3 that the months with a higher unemployment rate have a higher default (delinquency) risk. For Detroit and Las Vegas, when the unemployment rate is below the second quartile, its effect is positive which supports H3. However, when the unemployment rate is above the second quartile, its effect becomes insignificantly negative which conflict with H3.

For Miami, Tampa, and Las Vegas, the positive effect of the value of negative equity and the negative equity dummy supports the argument that when the value of the put option is “in the money,” households have more incentive to default (to be delinquent). However, the effect of the negative equity dummy is more significant compared with the effect of the value of negative

Table 28. Coefficients for Miami: Prepayment and default are dependent variables, the credit score only for default

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	the negative equity dummy and original loan-to-value are explanatory variables		negative equity and original loan-to-value are explanatory variables		the negative equity dummy is one of explanatory variables	
	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.578 (0.416)	--	-0.578 (0.416)	--	-0.578 (0.416)	--
Call option variables						
Call option part 1	4.932 (1.692)	--	4.932 (1.692)	--	4.931 (1.692)	--
Call option part 2	0.993 (1.765)	--	0.992 (1.765)	--	0.992 (1.765)	--
Call option part 3	8.381 (1.440)	--	8.384 (1.440)	--	8.381 (1.440)	--
Call option part 4	3.486 (0.935)	--	3.488 (0.935)	--	3.482 (0.935)	--
The unemployment rate variables						
The unemployment rate part 1	--	0.486 (0.132)	--	0.478 (0.132)	--	0.429 (0.132)
The unemployment rate part 2	--	0.135 (0.088)	--	0.115 (0.088)	--	0.121 (0.087)
Negative equity variables						
Negative equity part 1	--	--	--	0.061 (0.017)	--	--
Negative equity part 2	--	--	--	-0.004 (0.016)	--	--
Negative equity part 3	--	--	--	0.020 (0.010)	--	--
Negative equity part 4	--	--	--	0.002 (0.004)	--	--
The negative equity dummy	--	0.907 (0.136)	--	--	--	1.375 (0.121)
Original loan-to-value variables						
Original LTV part 1	--	3.250 (1.528)	--	3.872 (1.528)	--	--
Original LTV part 2	--	2.306 (1.924)	--	1.443 (1.949)	--	--
Original LTV part 3	--	5.332 (5.891)	--	5.216 (5.908)	--	--
Original LTV part 4	--	2.539 (1.022)	--	2.069 (1.040)	--	--
Original LTV part 5	--	14.002 (6.101)	--	12.967 (6.155)	--	--
Log loan size variables						
Log loan size part 1	0.644 (0.145)	0.630 (0.600)	0.644 (0.145)	0.720 (0.600)	0.644 (0.145)	0.709 (0.592)
Log loan size part 2	0.084 (0.269)	-0.390 (0.666)	0.084 (0.269)	-0.547 (0.677)	0.084 (0.269)	-0.002 (0.662)
Log loan size part 3	0.914 (0.283)	0.649 (0.530)	0.914 (0.283)	0.302 (0.548)	0.914 (0.283)	0.623 (0.524)
Log loan size part 4	0.060 (0.255)	0.326 (0.344)	0.060 (0.255)	0.041 (0.387)	0.060 (0.255)	0.323 (0.340)
Loan age variables						

Loan age part 1	0.125	0.168	0.125	0.172	0.125	0.166
	(0.014)	(0.047)	(0.014)	(0.047)	(0.014)	(0.047)
Loan age part 2	0.019	0.004	0.019	-0.005	0.019	-0.005
	(0.006)	(0.014)	(0.006)	(0.015)	(0.006)	(0.014)
Loan age part 3	-0.001	-0.022	-0.001	-0.021	-0.001	-0.022
	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.007)
Loan age part 4	0.003	-0.018	0.003	-0.016	0.003	-0.008
	(0.002)	(0.006)	(0.002)	(0.005)	(0.002)	(0.005)
The credit score variables						
The credit score part 1	--	-9.114	--	-9.129	--	-7.377
		(2.710)		(2.723)		(2.657)
The credit score part 2	--	-5.223	--	-5.018	--	-5.890
		(4.596)		(4.630)		(4.567)
The credit score part 3	--	-7.645	--	-7.875	--	-7.723
		(5.042)		(5.064)		(4.998)
The credit score part 4	--	-20.745	--	-21.235	--	-20.816
		(7.021)		(7.021)		(6.992)
The debt-to-income ratio variables						
DTI part 1	-0.044	2.753	-0.044	2.631	-0.044	3.533
	(0.738)	(2.237)	(0.738)	(2.225)	(0.738)	(2.253)
DTI part 2	0.265	-0.715	0.265	-0.620	0.266	-0.350
	(0.949)	(1.992)	(0.949)	(1.994)	(0.949)	(2.011)
DTI part 3	-1.812	1.516	-1.812	1.423	-1.812	1.242
	(1.204)	(2.272)	(1.204)	(2.287)	(1.204)	(2.283)
DTI part 4	0.999	-2.069	0.999	-1.790	0.999	-2.023
	(0.722)	(1.213)	(0.722)	(1.221)	(0.722)	(1.209)
The dummy for the number of units	-0.451	-0.040	-0.451	-0.040	-0.451	-0.107
	(0.213)	(0.388)	(0.213)	(0.389)	(0.213)	(0.388)
Time period variables						
Time period part 1	0.031	0.014	0.031	0.015	0.031	0.007
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Time period part 2	-0.037	0.047	-0.037	0.052	-0.037	0.048
	(0.002)	(0.025)	(0.002)	(0.025)	(0.002)	(0.025)
Time period part 3	0.014	-0.014	0.014	-0.016	0.014	-0.018
	(0.003)	(0.011)	(0.003)	(0.011)	(0.003)	(0.011)

Number of mortgages 3224

Table 29. Odds ratio for Miami: Prepayment and default are dependent variables, the credit score only for default

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	the negative equity dummy and original loan-to-value are explanatory variables		negative equity and original loan-to-value are explanatory variables		the negative equity dummy is one of explanatory variables	
	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	0.561	--	0.561	--	0.561	--
	(0.233)		(0.233)		(0.233)	
Call option variables						
Call option part 1	138.593	--	138.669	--	138.530	--
	(234.520)		(234.651)		(234.414)	
Call option part 2	2.700	--	2.697	--	2.697	--
	(4.765)		(4.761)		(4.759)	
Call option part 3	4,364.551	--	4,375.890	--	4,363.658	--
	(6,286.386)		(6,302.792)		(6,285.026)	

Call option part 4	32.663 (30.548)	--	32.714 (30.597)	--	32.510 (30.402)	--
The unemployment rate variables						
The unemployment rate part 1	--	1.626 (0.215)	--	1.613 (0.213)	--	1.536 (0.204)
The unemployment rate part 2	--	1.144 (0.100)	--	1.122 (0.099)	--	1.129 (0.098)
Negative equity variables						
Negative equity part 1	--	--	--	1.063 (0.018)	--	--
Negative equity part 2	--	--	--	0.996 (0.016)	--	--
Negative equity part 3	--	--	--	1.020 (0.010)	--	--
Negative equity part 4	--	--	--	1.002 (0.004)	--	--
The negative equity dummy	--	2.477 (0.338)	--	--	--	3.955 (0.480)
Original loan-to-value variables						
Original LTV part 1	--	25.801 (39.427)	--	48.028 (73.405)	--	--
Original LTV part 2	--	10.038 (19.314)	--	4.235 (8.251)	--	--
Original LTV part 3	--	206.814 (1,218.385)	--	184.241 (1,088.523)	--	--
Original LTV part 4	--	12.670 (12.943)	--	7.918 (8.238)	--	--
Original LTV part 5	--	1204641.781 (7349619.485)	--	428,193.827 (2635361.371)	--	--
Log loan size variables						
Log loan size part 1	1.905 (0.277)	1.877 (1.126)	1.905 (0.277)	2.055 (1.234)	1.904 (0.277)	2.032 (1.203)
Log loan size part 2	1.087 (0.293)	0.677 (0.451)	1.087 (0.293)	0.578 (0.392)	1.088 (0.293)	0.998 (0.661)
Log loan size part 3	2.494 (0.706)	1.913 (1.014)	2.494 (0.706)	1.352 (0.740)	2.493 (0.706)	1.865 (0.977)
Log loan size part 4	1.062 (0.271)	1.386 (0.477)	1.062 (0.271)	1.042 (0.403)	1.062 (0.271)	1.381 (0.469)
Loan age variables						
Loan age part 1	1.133 (0.016)	1.183 (0.056)	1.133 (0.016)	1.188 (0.056)	1.133 (0.016)	1.181 (0.056)
Loan age part 2	1.019 (0.007)	1.004 (0.014)	1.019 (0.007)	0.995 (0.015)	1.019 (0.007)	0.995 (0.014)
Loan age part 3	0.999 (0.004)	0.979 (0.007)	0.999 (0.004)	0.979 (0.007)	0.999 (0.004)	0.978 (0.007)
Loan age part 4	1.003 (0.002)	0.983 (0.005)	1.003 (0.002)	0.985 (0.005)	1.003 (0.002)	0.992 (0.005)
The credit score variables						
The credit score part 1	--	0.000 (0.000)	--	0.000 (0.000)	--	0.001 (0.002)
The credit score part 2	--	0.005 (0.025)	--	0.007 (0.031)	--	0.003 (0.013)
The credit score part 3	--	0.000 (0.002)	--	0.000 (0.002)	--	0.000 (0.002)
The credit score part 4	--	0.000 (0.000)	--	0.000 (0.000)	--	0.000 (0.000)
The debt-to-income ratio variables						
DTI part 1	0.957 (0.706)	15.694 (35.114)	0.957 (0.706)	13.881 (30.881)	0.957 (0.706)	34.226 (77.105)
DTI part 2	1.304	0.489	1.304	0.538	1.304	0.705

	(1.237)	(0.975)	(1.238)	(1.073)	(1.238)	(1.417)
DTI part 3	0.163	4.555	0.163	4.148	0.163	3.461
	(0.197)	(10.350)	(0.197)	(9.489)	(0.197)	(7.900)
DTI part 4	2.715	0.126	2.715	0.167	2.717	0.132
	(1.959)	(0.153)	(1.959)	(0.204)	(1.960)	(0.160)
The dummy for the number of units	0.637	0.961	0.637	0.961	0.637	0.898
	(0.136)	(0.373)	(0.136)	(0.373)	(0.136)	(0.348)
Time period variables						
Time period part 1	1.032	1.014	1.032	1.015	1.032	1.007
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)
Time period part 2	0.963	1.048	0.963	1.054	0.963	1.049
	(0.002)	(0.026)	(0.002)	(0.026)	(0.002)	(0.026)
Time period part 3	1.014	0.986	1.014	0.984	1.014	0.983
	(0.003)	(0.011)	(0.003)	(0.011)	(0.003)	(0.011)

Number of mortgages 3224

Table 30. Coefficients for Miami: Prepayment and 90-days-delinquency are dependent variables, the credit score only for 90-days-delinquency

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	the negative equity dummy and original loan-to-value are explanatory variables		the negative equity dummy is one of explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	-0.567	--	-0.567	--	-0.567	--
	(0.416)		(0.416)		(0.416)	
Call option variables						
Call option part 1	5.255	--	5.254	--	5.255	--
	(1.740)		(1.740)		(1.740)	
Call option part 2	1.038	--	1.037	--	1.038	--
	(1.804)		(1.804)		(1.804)	
Call option part 3	8.678	--	8.677	--	8.679	--
	(1.454)		(1.454)		(1.454)	
Call option part 4	3.673	--	3.669	--	3.675	--
	(0.929)		(0.929)		(0.929)	
The unemployment rate variables						
The unemployment rate part 1	--	0.588	--	0.540	--	0.583
		(0.108)		(0.108)		(0.108)
The unemployment rate part 2	--	0.126	--	0.110	--	0.109
		(0.081)		(0.081)		(0.081)
Negative equity variables						
Negative equity part 1	--	--	--	--	--	0.053
						(0.016)
Negative equity part 2	--	--	--	--	--	-0.006
						(0.016)
Negative equity part 3	--	--	--	--	--	0.025
						(0.010)
Negative equity part 4	--	--	--	--	--	-0.003
						(0.004)
The negative equity dummy	--	0.802	--	1.219	--	--
		(0.123)		(0.109)		
Original loan-to-value						

variables						
Original LTV part 1	--	2.252	--	--	--	2.761
		(1.258)				(1.261)
Original LTV part 2	--	2.616	--	--	--	1.924
		(1.738)				(1.755)
Original LTV part 3	--	4.775	--	--	--	4.873
		(5.427)				(5.445)
Original LTV part 4	--	2.306	--	--	--	2.088
		(0.958)				(0.971)
Original LTV part 5	--	10.537	--	--	--	9.110
		(6.095)				(6.134)
Log loan size variables						
Log loan size part 1	0.646	0.559	0.646	0.594	0.646	0.642
	(0.147)	(0.553)	(0.147)	(0.545)	(0.147)	(0.554)
Log loan size part 2	0.089	0.164	0.090	0.515	0.089	0.013
	(0.269)	(0.624)	(0.269)	(0.620)	(0.269)	(0.633)
Log loan size part 3	0.923	0.276	0.923	0.283	0.923	-0.050
	(0.282)	(0.487)	(0.282)	(0.482)	(0.282)	(0.500)
Log loan size part 4	0.068	0.430	0.068	0.424	0.068	0.325
	(0.253)	(0.317)	(0.253)	(0.313)	(0.253)	(0.345)
Loan age variables						
Loan age part 1	0.134	0.092	0.134	0.088	0.134	0.096
	(0.016)	(0.034)	(0.016)	(0.033)	(0.016)	(0.033)
Loan age part 2	0.022	0.002	0.022	-0.006	0.022	-0.005
	(0.006)	(0.012)	(0.006)	(0.012)	(0.006)	(0.012)
Loan age part 3	-0.001	-0.026	-0.001	-0.024	-0.001	-0.024
	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.007)
Loan age part 4	0.003	-0.014	0.003	-0.006	0.003	-0.013
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)
The credit score variables						
The credit score part 1	--	-10.230	--	-8.606	--	-9.997
		(2.397)		(2.350)		(2.411)
The credit score part 2	--	-5.342	--	-5.992	--	-5.550
		(4.127)		(4.104)		(4.162)
The credit score part 3	--	-11.579	--	-11.670	--	-11.418
		(4.914)		(4.882)		(4.930)
The credit score part 4	--	-19.880	--	-19.826	--	-20.752
		(6.819)		(6.789)		(6.822)
The debt-to-income ratio variables						
DTI part 1	0.008	3.494	0.008	4.181	0.008	3.381
	(0.746)	(2.192)	(0.746)	(2.208)	(0.746)	(2.186)
DTI part 2	0.160	-0.514	0.160	-0.062	0.160	-0.361
	(1.064)	(2.103)	(1.064)	(2.125)	(1.064)	(2.105)
DTI part 3	-1.403	2.386	-1.403	2.161	-1.404	2.297
	(1.075)	(1.880)	(1.075)	(1.888)	(1.075)	(1.894)
DTI part 4	0.823	-2.646	0.824	-2.713	0.823	-2.332
	(0.719)	(1.135)	(0.719)	(1.131)	(0.719)	(1.141)
The dummy for the number of units	-0.458	-0.151	-0.458	-0.219	-0.458	-0.171
	(0.213)	(0.387)	(0.213)	(0.387)	(0.213)	(0.387)
Time period variables						
Time period part 1	0.030	0.022	0.030	0.015	0.030	0.023
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Time period part 2	-0.037	0.015	-0.037	0.016	-0.037	0.020
	(0.002)	(0.024)	(0.002)	(0.024)	(0.002)	(0.024)
Time period part 3	0.014	-0.021	0.014	-0.025	0.014	-0.023
	(0.003)	(0.011)	(0.003)	(0.011)	(0.003)	(0.011)

Number of mortgages 3211

Table 31. Odds ratios for Miami: Prepayment and 90-days-delinquency are dependent variables, the credit score only for 90-days-delinquency

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	the negative equity dummy and original loan-to-value are explanatory variables		the negative equity dummy is one of explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	0.567 (0.236)	--	0.567 (0.236)	--	0.567 (0.236)	--
Call option variables						
Call option part 1	191.452 (333.035)	--	191.372 (332.897)	--	191.551 (333.210)	--
Call option part 2	2.823 (5.092)	--	2.820 (5.086)	--	2.822 (5.090)	--
Call option part 3	5,871.224 (8,537.001)	--	5,867.609 (8,531.650)	--	5,880.645 (8,550.805)	--
Call option part 4	39.384 (36.580)	--	39.195 (36.401)	--	39.458 (36.651)	--
The unemployment rate variables						
The unemployment rate part 1	--	1.800 (0.194)	--	1.717 (0.185)	--	1.792 (0.193)
The unemployment rate part 2	--	1.135 (0.092)	--	1.116 (0.090)	--	1.115 (0.091)
Negative equity variables						
Negative equity part 1	--	--	--	--	--	1.054 (0.017)
Negative equity part 2	--	--	--	--	--	0.994 (0.016)
Negative equity part 3	--	--	--	--	--	1.026 (0.010)
Negative equity part 4	--	--	--	--	--	0.997 (0.004)
The negative equity dummy	--	2.229 (0.274)	--	3.385 (0.369)	--	--
Original loan-to-value variables						
Original LTV part 1	--	9.508 (11.956)	--	--	--	15.821 (19.950)
Original LTV part 2	--	13.675 (23.761)	--	--	--	6.847 (12.018)
Original LTV part 3	--	118.507 (643.129)	--	--	--	130.738 (711.834)
Original LTV part 4	--	10.032 (9.608)	--	--	--	8.073 (7.842)
Original LTV part 5	--	37,692.552 (229,744.910)	--	--	--	9,044.227 (55,481.191)
Log loan size variables						
Log loan size part 1	1.909 (0.281)	1.749 (0.968)	1.908 (0.281)	1.811 (0.988)	1.909 (0.281)	1.901 (1.053)
Log loan size part 2	1.094 (0.294)	1.178 (0.735)	1.094 (0.294)	1.674 (1.038)	1.094 (0.294)	1.013 (0.641)
Log loan size part 3	2.517 (0.709)	1.318 (0.642)	2.517 (0.709)	1.328 (0.640)	2.517 (0.709)	0.952 (0.476)
Log loan size part 4	1.071 (0.271)	1.537 (0.487)	1.071 (0.271)	1.527 (0.478)	1.071 (0.271)	1.385 (0.478)

Loan age variables						
Loan age part 1	1.144	1.097	1.144	1.092	1.144	1.101
	(0.018)	(0.037)	(0.018)	(0.037)	(0.018)	(0.037)
Loan age part 2	1.022	1.002	1.022	0.994	1.022	0.995
	(0.006)	(0.012)	(0.006)	(0.012)	(0.006)	(0.012)
Loan age part 3	0.999	0.975	0.999	0.976	0.999	0.976
	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.007)
Loan age part 4	1.003	0.986	1.003	0.994	1.003	0.987
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)
The credit score variables						
The credit score part 1	--	0.000	--	0.000	--	0.000
		(0.000)		(0.000)		(0.000)
The credit score part 2	--	0.005	--	0.002	--	0.004
		(0.020)		(0.010)		(0.016)
The credit score part 3	--	0.000	--	0.000	--	0.000
		(0.000)		(0.000)		(0.000)
The credit score part 4	--	0.000	--	0.000	--	0.000
		(0.000)		(0.000)		(0.000)
The debt-to-income ratio variables						
DTI part 1	1.008	32.914	1.008	65.453	1.008	29.409
	(0.752)	(72.155)	(0.751)	(144.506)	(0.752)	(64.284)
DTI part 2	1.173	0.598	1.174	0.940	1.174	0.697
	(1.249)	(1.257)	(1.249)	(1.996)	(1.249)	(1.467)
DTI part 3	0.246	10.867	0.246	8.676	0.246	9.949
	(0.264)	(20.433)	(0.264)	(16.380)	(0.264)	(18.843)
DTI part 4	2.278	0.071	2.279	0.066	2.277	0.097
	(1.638)	(0.080)	(1.639)	(0.075)	(1.638)	(0.111)
The dummy for the number of units	0.633	0.860	0.633	0.803	0.633	0.843
	(0.135)	(0.333)	(0.135)	(0.310)	(0.135)	(0.326)
Time period variables						
Time period part 1	1.030	1.022	1.030	1.015	1.030	1.023
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Time period part 2	0.964	1.015	0.964	1.016	0.964	1.020
	(0.002)	(0.024)	(0.002)	(0.024)	(0.002)	(0.024)
Time period part 3	1.014	0.979	1.014	0.975	1.014	0.977
	(0.003)	(0.011)	(0.003)	(0.011)	(0.003)	(0.011)

Number of mortgages 3211

Table 32. Results comparison in Miami

Prepayment and Default are Dependent Variables

		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and default		The credit score for default		The credit score for prepayment and default		The credit score for default	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	original LTV, the negative equity dummy	0.9282	-11694.2	0.9282	-11680.7	0.9274	-11443.9	0.9274	-11651.6
2	negative equity, original LTV	0.9282	-11700.7	0.9282	-11687.2	0.9274	-11438.1	0.9274	-11657.9

Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
3	the negative equity dummy	0.928	-11701.2	0.928	-11687.7	0.9272	-11470.7	0.9272	-11658.3

		Prepayment and 90-days-delinquency are Dependent Variables							
		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and 90-days-delinquency		The credit score for 90-days-delinquency		The credit score for prepayment and 90-days-delinquency		The credit score for 90-days-delinquency	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	original LTV, the negative equity dummy	0.9248	-11956.8	0.9247	-11942.8	0.9249	-11779.8	0.9249	-11765.5
2	the negative equity dummy	0.9246	-11962.3	0.9246	-11948.3	0.9248	-11785.9	0.9248	-11771.6
3	negative equity, original LTV	0.9248	-11963.7	0.9248	-11949.7	0.925	-11786.7	0.925	-11772.4

Table 33. Coefficients for Tampa: Prepayment and default are dependent variables, the credit score for both the prepayment and the default

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	the negative equity dummy is one of explanatory variables		the negative equity dummy and original loan-to-value are explanatory variables		negative equity and the negative equity dummy are explanatory variables	
	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.723 (0.338)	--	-0.723 (0.338)	--	-0.723 (0.338)	--
Call option variables						
Call option part 1	2.233 (1.145)	--	2.233 (1.145)	--	2.233 (1.145)	--
Call option part 2	7.040 (1.255)	--	7.041 (1.255)	--	7.040 (1.255)	--
Call option part 3	7.647 (0.977)	--	7.647 (0.977)	--	7.649 (0.977)	--
Call option part 4	1.858 (0.656)	--	1.858 (0.656)	--	1.859 (0.656)	--
The unemployment rate variables						
The unemployment rate part 1	--	0.591 (0.179)	--	0.599 (0.179)	--	0.624 (0.180)
The unemployment rate part 2	--	0.013 (0.060)	--	0.004 (0.059)	--	0.012 (0.060)
Negative equity variables						
Negative equity part 1	--	--	--	--	--	0.030 (0.033)
Negative equity part 2	--	--	--	--	--	0.009 (0.020)

Negative equity part 3	--	--	--	--	--	0.015
						(0.013)
Negative equity part 4	--	--	--	--	--	0.006
						(0.005)
The negative equity dummy	--	1.234	--	0.974		0.837
		(0.114)		(0.125)		(0.264)
Original loan-to-value variables						
Original LTV part 1	--	--	--	1.347	--	--
				(1.118)		
Original LTV part 2	--	--	--	2.003	--	--
				(1.216)		
Original LTV part 3	--	--	--	3.266	--	--
				(0.872)		
Original LTV part 4	--	--	--	-4.650	--	--
				(6.120)		
Original LTV part 5	--	--	--	--	--	--
Log loan size variables						
Log loan size part 1	0.533	0.420	0.533	0.381	0.533	0.447
	(0.104)	(0.414)	(0.104)	(0.419)	(0.104)	(0.414)
Log loan size part 2	0.849	-0.046	0.849	-0.031	0.849	-0.133
	(0.177)	(0.515)	(0.177)	(0.514)	(0.177)	(0.518)
Log loan size part 3	0.082	-0.066	0.082	-0.156	0.082	-0.405
	(0.198)	(0.474)	(0.198)	(0.479)	(0.198)	(0.496)
Log loan size part 4	0.687	-0.055	0.687	-0.096	0.687	-0.330
	(0.136)	(0.303)	(0.136)	(0.305)	(0.136)	(0.343)
Loan age variables						
Loan age part 1	0.123	0.168	0.123	0.167	0.123	0.170
	(0.010)	(0.058)	(0.010)	(0.058)	(0.010)	(0.058)
Loan age part 2	0.010	0.030	0.010	0.035	0.010	0.027
	(0.005)	(0.017)	(0.005)	(0.017)	(0.005)	(0.017)
Loan age part 3	0.004	-0.012	0.004	-0.012	0.004	-0.015
	(0.003)	(0.007)	(0.003)	(0.007)	(0.003)	(0.007)
Loan age part 4	0.009	-0.002	0.009	-0.007	0.009	-0.000
	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)
The credit score variables						
The credit score part 1	1.248	-7.010	1.247	-7.529	1.248	-7.262
	(1.100)	(2.216)	(1.100)	(2.194)	(1.100)	(2.229)
The credit score part 2	1.492	-7.116	1.492	-6.467	1.492	-6.349
	(1.464)	(3.712)	(1.464)	(3.696)	(1.464)	(3.719)
The credit score part 3	2.000	-6.078	2.000	-6.451	2.000	-6.554
	(1.552)	(4.687)	(1.552)	(4.682)	(1.552)	(4.691)
The credit score part 4	5.670	-25.519	5.670	-24.291	5.669	-24.578
	(2.102)	(8.400)	(2.102)	(8.380)	(2.102)	(8.385)
The debt-to-income ratio variables						
DTI part 1	-0.683	-3.425	-0.683	-3.753	-0.683	-3.373
	(0.511)	(1.762)	(0.511)	(1.774)	(0.511)	(1.759)
DTI part 2	0.191	4.606	0.191	4.522	0.191	4.656
	(0.733)	(2.184)	(0.733)	(2.181)	(0.733)	(2.187)
DTI part 3	-0.593	1.014	-0.592	0.662	-0.592	0.837
	(0.832)	(1.975)	(0.832)	(1.977)	(0.832)	(1.977)
DTI part 4	0.175	0.616	0.175	0.685	0.175	0.518
	(0.536)	(1.080)	(0.536)	(1.088)	(0.536)	(1.079)
The dummy for the number of units	-0.626	-0.563	-0.626	-0.532	-0.626	-0.511
	(0.212)	(0.581)	(0.212)	(0.581)	(0.212)	(0.581)
Time period variables						
Time period part 1	0.021	0.008	0.021	0.010	0.021	0.010
	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)	(0.005)
Time period part 2	-0.037	0.042	-0.037	0.054	-0.037	0.035
	(0.001)	(0.036)	(0.001)	(0.036)	(0.001)	(0.036)
Time period part 3	0.012	0.007	0.010	0.008	0.010	0.007
	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)

Number of mortgages 6129

Table 34. Odds ratio for Tampa: Prepayment and default are dependent variables, the credit score for both the prepayment and the default

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	the negative equity dummy is one of explanatory variables		the negative equity dummy and original loan-to-value are explanatory variables		negative equity and the negative equity dummy are explanatory variables	
	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	0.485	--	0.485	--	0.485	--
	(0.164)		(0.164)		(0.164)	
Call option variables						
Call option part 1	9.327	--	9.326	--	9.329	--
	(10.678)		(10.677)		(10.682)	
Call option part 2	1,141.531	--	1,142.131	--	1,140.857	--
	(1,432.569)		(1,433.323)		(1,431.724)	
Call option part 3	2,094.932	--	2,094.291	--	2,098.459	--
	(2,046.269)		(2,045.644)		(2,049.712)	
Call option part 4	6.409	--	6.412	--	6.417	--
	(4.203)		(4.205)		(4.208)	
The unemployment rate variables						
The unemployment rate part 1	--	1.806	--	1.820	--	1.867
		(0.323)		(0.326)		(0.336)
The unemployment rate part 2	--	1.013	--	1.004	--	1.012
		(0.060)		(0.060)		(0.061)
Negative equity variables						
Negative equity part 1	--	--	--	--	--	1.030
						(0.034)
Negative equity part 2	--	--	--	--	--	1.009
						(0.020)
Negative equity part 3	--	--	--	--	--	1.015
						(0.014)
Negative equity part 4	--	--	--	--	--	1.006
						(0.006)
The negative equity dummy	--	3.435	--	2.648		2.310
		(0.390)		(0.330)		(0.609)
Original loan-to-value variables						
Original LTV part 1	--	--	--	3.844	--	--
				(4.297)		
Original LTV part 2	--	--	--	7.408	--	--
				(9.005)		
Original LTV part 3	--	--	--	26.204	--	--
				(22.857)		
Original LTV part 4	--	--	--	0.010	--	--
				(0.059)		
Original LTV part 5	--	--	--	--	--	--
Log loan size variables						
Log loan size part 1	1.704	1.522	1.704	1.463	1.704	1.563
	(0.178)	(0.631)	(0.178)	(0.614)	(0.178)	(0.647)
Log loan size part 2	2.338	0.955	2.338	0.970	2.338	0.876

	(0.414)	(0.492)	(0.414)	(0.499)	(0.414)	(0.454)
Log loan size part 3	1.086	0.936	1.086	0.856	1.086	0.667
	(0.215)	(0.444)	(0.215)	(0.410)	(0.215)	(0.331)
Log loan size part 4	1.987	0.946	1.987	0.908	1.987	0.719
	(0.270)	(0.287)	(0.270)	(0.277)	(0.270)	(0.247)
Loan age variables						
Loan age part 1	1.130	1.183	1.130	1.182	1.130	1.185
	(0.011)	(0.069)	(0.011)	(0.069)	(0.011)	(0.069)
Loan age part 2	1.010	1.031	1.010	1.036	1.010	1.027
	(0.005)	(0.017)	(0.005)	(0.017)	(0.005)	(0.017)
Loan age part 3	1.004	0.988	1.004	0.988	1.004	0.985
	(0.003)	(0.007)	(0.003)	(0.007)	(0.003)	(0.007)
Loan age part 4	1.009	0.998	1.009	0.993	1.009	1.000
	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)
The credit score variables						
The credit score part 1	3.482	0.001	3.480	0.001	3.482	0.001
	(3.831)	(0.002)	(3.829)	(0.001)	(3.831)	(0.002)
The credit score part 2	4.447	0.001	4.447	0.002	4.447	0.002
	(6.513)	(0.003)	(6.513)	(0.006)	(6.513)	(0.007)
The credit score part 3	7.387	0.002	7.388	0.002	7.389	0.001
	(11.468)	(0.011)	(11.469)	(0.007)	(11.471)	(0.007)
The credit score part 4	290.016	0.000	289.981	0.000	289.843	0.000
	(609.607)	(0.000)	(609.536)	(0.000)	(609.245)	(0.000)
The debt-to-income ratio variables						
DTI part 1	0.505	0.033	0.505	0.023	0.505	0.034
	(0.258)	(0.057)	(0.258)	(0.042)	(0.258)	(0.060)
DTI part 2	1.210	100.131	1.210	91.988	1.210	105.232
	(0.887)	(218.661)	(0.887)	(200.625)	(0.887)	(230.139)
DTI part 3	0.553	2.757	0.553	1.938	0.553	2.310
	(0.460)	(5.445)	(0.460)	(3.831)	(0.460)	(4.566)
DTI part 4	1.191	1.851	1.191	1.984	1.191	1.678
	(0.639)	(2.000)	(0.639)	(2.159)	(0.639)	(1.811)
The dummy for the number of units	0.535	0.570	0.534	0.587	0.534	0.600
	(0.113)	(0.331)	(0.113)	(0.342)	(0.113)	(0.349)
Time period variables						
Time period part 1	1.021	1.008	1.021	1.010	1.021	1.010
	(0.003)	(0.005)	(0.003)	(0.006)	(0.003)	(0.006)
Time period part 2	0.963	1.043	0.963	1.055	0.963	1.035
	(0.001)	(0.037)	(0.001)	(0.038)	(0.001)	(0.037)
Time period part 3	1.010	1.007	1.010	1.008	1.010	1.007
	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)

Number of mortgages 6129

Table 35. Coefficients for Tampa: Prepayment and 90-days-delinquency are dependent variables, the credit score for both the prepayment and the 90-days-delinquency

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	the negative equity dummy is one of explanatory variables		the negative equity dummy and original loan-to-value are explanatory variables		negative equity is one of explanatory variables	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	-0.721	--	-0.721	--	-0.722	--

	(0.338)		(0.338)		(0.338)	
Call option variables						
Call option part 1	2.359	--	2.359	--	2.359	--
	(1.161)		(1.161)		(1.161)	
Call option part 2	6.919	--	6.920	--	6.919	--
	(1.278)		(1.278)		(1.278)	
Call option part 3	7.890	--	7.890	--	7.891	--
	(0.982)		(0.982)		(0.982)	
Call option part 4	2.061	--	2.061	--	2.061	--
	(0.653)		(0.653)		(0.653)	
The unemployment rate variables						
The unemployment rate part 1	--	0.436	--	0.446	--	0.472
		(0.127)		(0.128)		(0.128)
The unemployment rate part 2	--	0.028	--	0.021	--	0.027
		(0.055)		(0.054)		(0.055)
Negative equity variables						
Negative equity part 1	--	--	--		--	0.096
						(0.017)
Negative equity part 2	--	--	--		--	0.007
						(0.018)
Negative equity part 3	--	--	--		--	0.008
						(0.013)
Negative equity part 4	--	--	--		--	0.000
						(0.006)
The negative equity dummy	--	1.022	--	0.766	--	--
		(0.101)		(0.111)		
Original loan-to-value variables						
Original LTV part 1	--	--	--	2.353	--	--
				(1.001)		
Original LTV part 2	--	--	--	1.045	--	--
				(1.074)		
Original LTV part 3	--	--	--	3.062	--	--
				(0.811)		
Original LTV part 4	--	--	--	-5.963	--	--
				(5.901)		
Original LTV part 5	--	--	--	--	--	--
Log loan size variables						
Log loan size part 1	0.543	0.424	0.543	0.342	0.543	0.426
	(0.105)	(0.353)	(0.105)	(0.360)	(0.105)	(0.351)
Log loan size part 2	0.865	-0.221	0.865	-0.246	0.865	-0.280
	(0.177)	(0.456)	(0.177)	(0.456)	(0.177)	(0.457)
Log loan size part 3	0.077	-0.195	0.077	-0.263	0.077	-0.444
	(0.198)	(0.436)	(0.198)	(0.440)	(0.198)	(0.454)
Log loan size part 4	0.684	-0.036	0.684	-0.077	0.684	-0.128
	(0.136)	(0.279)	(0.136)	(0.281)	(0.136)	(0.306)
Loan age variables						
Loan age part 1	0.133	0.133	0.133	0.132	0.133	0.137
	(0.011)	(0.044)	(0.011)	(0.044)	(0.011)	(0.044)
Loan age part 2	0.015	0.029	0.015	0.034	0.015	0.028
	(0.005)	(0.014)	(0.005)	(0.014)	(0.005)	(0.014)
Loan age part 3	0.003	-0.017	0.003	-0.017	0.003	-0.018
	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)	(0.006)
Loan age part 4	0.009	-0.004	0.009	-0.009	0.009	-0.004
	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)
The credit score variables						
The credit score part 1	1.152	-8.766	1.152	-9.117	1.153	-8.928
	(1.104)	(1.891)	(1.104)	(1.858)	(1.104)	(1.890)
The credit score part 2	1.311	-9.215	1.311	-8.762	1.311	-8.996
	(1.437)	(3.333)	(1.437)	(3.312)	(1.437)	(3.335)
The credit score part 3	1.902	-7.200	1.902	-7.495	1.902	-7.333

	(1.594)	(4.541)	(1.594)	(4.537)	(1.594)	(4.543)
The credit score part 4	5.488	-27.576	5.488	-26.365	5.488	-27.141
	(2.108)	(8.162)	(2.108)	(8.142)	(2.108)	(8.154)
The debt-to-income ratio variables						
DTI part 1	-0.733	-2.521	-0.733	-2.882	-0.733	-2.413
	(0.512)	(1.683)	(0.512)	(1.697)	(0.512)	(1.679)
DTI part 2	0.209	5.062	0.209	4.988	0.209	5.185
	(0.735)	(2.013)	(0.735)	(2.013)	(0.735)	(2.015)
DTI part 3	-0.520	1.441	-0.520	1.104	-0.520	1.355
	(0.834)	(1.792)	(0.834)	(1.794)	(0.834)	(1.792)
DTI part 4	0.106	0.522	0.106	0.662	0.106	0.477
	(0.538)	(0.976)	(0.538)	(0.983)	(0.538)	(0.975)
The dummy for the number of units	-0.626	-0.424	-0.626	-0.380	-0.626	-0.366
	(0.212)	(0.504)	(0.212)	(0.504)	(0.212)	(0.504)
Time period variables						
Time period part 1	0.021	0.010	0.021	0.013	0.021	0.012
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Time period part 2	-0.037	0.035	-0.037	0.046	-0.037	0.031
	(0.001)	(0.031)	(0.001)	(0.031)	(0.001)	(0.031)
Time period part 3	0.010	0.001	0.010	0.002	0.010	0.001
	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)

Number of mortgages 6112

Table 36. Odds ratios for Tampa: Prepayment and 90-days-delinquency are dependent variables, the credit score for both the prepayment and the 90-days-delinquency

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	the negative equity dummy is one of explanatory variables		the negative equity dummy and original loan-to-value are explanatory variables		negative equity is one of explanatory variables	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	0.486	--	0.486	--	0.486	--
	(0.164)		(0.164)		(0.164)	
Call option variables						
Call option part 1	10.578	--	10.577	--	10.580	--
	(12.277)		(12.276)		(12.279)	
Call option part 2	1,011.708	--	1,012.267	--	1,011.313	--
	(1,293.275)		(1,293.989)		(1,292.770)	
Call option part 3	2,670.345	--	2,670.144	--	2,673.736	--
	(2,621.138)		(2,620.943)		(2,624.465)	
Call option part 4	7.854	--	7.855	--	7.858	--
	(5.127)		(5.128)		(5.130)	
The unemployment rate variables						
The unemployment rate part 1	--	1.546	--	1.562	--	1.603
		(0.197)		(0.199)		(0.205)
The unemployment rate part 2	--	1.029	--	1.022	--	1.028
		(0.056)		(0.055)		(0.056)
Negative equity variables						
Negative equity part 1	--	--	--		--	1.101
						(0.018)
Negative equity part 2	--	--	--		--	1.007

						(0.018)
Negative equity part 3	--	--	--		--	1.009
						(0.013)
Negative equity part 4	--	--	--		--	1.000
						(0.006)
The negative equity dummy	--	2.780	--	2.152	--	--
		(0.280)		(0.238)		
Original loan-to-value variables						
Original LTV part 1	--	--	--	10.515	--	--
				(10.526)		
Original LTV part 2	--	--	--	2.843	--	--
				(3.054)		
Original LTV part 3	--	--	--	21.372	--	--
				(17.328)		
Original LTV part 4	--	--	--	0.003	--	--
				(0.015)		
Original LTV part 5	--	--	--	--	--	--
Log loan size variables						
Log loan size part 1	1.721	1.528	1.721	1.407	1.721	1.532
	(0.181)	(0.539)	(0.181)	(0.506)	(0.181)	(0.538)
Log loan size part 2	2.376	0.802	2.376	0.782	2.376	0.756
	(0.422)	(0.365)	(0.422)	(0.356)	(0.422)	(0.345)
Log loan size part 3	1.080	0.822	1.080	0.769	1.080	0.642
	(0.214)	(0.359)	(0.214)	(0.338)	(0.214)	(0.291)
Log loan size part 4	1.981	0.964	1.981	0.925	1.981	0.880
	(0.270)	(0.269)	(0.270)	(0.260)	(0.270)	(0.269)
Loan age variables						
Loan age part 1	1.143	1.143	1.143	1.141	1.143	1.147
	(0.013)	(0.050)	(0.013)	(0.050)	(0.013)	(0.051)
Loan age part 2	1.015	1.030	1.015	1.034	1.015	1.028
	(0.005)	(0.014)	(0.005)	(0.014)	(0.005)	(0.014)
Loan age part 3	1.003	0.983	1.003	0.983	1.003	0.982
	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)	(0.006)
Loan age part 4	1.009	0.996	1.009	0.991	1.009	0.996
	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)
The credit score variables						
The credit score part 1	3.166	0.000	3.164	0.000	3.166	0.000
	(3.495)	(0.000)	(3.493)	(0.000)	(3.496)	(0.000)
The credit score part 2	3.710	0.000	3.709	0.000	3.709	0.000
	(5.330)	(0.000)	(5.330)	(0.001)	(5.329)	(0.000)
The credit score part 3	6.697	0.001	6.697	0.001	6.698	0.001
	(10.671)	(0.003)	(10.671)	(0.003)	(10.674)	(0.003)
The credit score part 4	241.788	0.000	241.772	0.000	241.696	0.000
	(509.773)	(0.000)	(509.744)	(0.000)	(509.580)	(0.000)
The debt-to-income ratio variables						
DTI part 1	0.480	0.080	0.480	0.056	0.480	0.090
	(0.246)	(0.135)	(0.246)	(0.095)	(0.246)	(0.150)
DTI part 2	1.233	157.962	1.233	146.624	1.233	178.610
	(0.906)	(318.038)	(0.906)	(295.208)	(0.906)	(359.864)
DTI part 3	0.594	4.225	0.594	3.015	0.594	3.876
	(0.496)	(7.571)	(0.496)	(5.407)	(0.496)	(6.947)
DTI part 4	1.112	1.685	1.112	1.938	1.112	1.611
	(0.599)	(1.645)	(0.598)	(1.906)	(0.599)	(1.570)
The dummy for the number of units	0.535	0.655	0.535	0.684	0.535	0.694
	(0.113)	(0.330)	(0.113)	(0.345)	(0.113)	(0.349)
Time period variables						
Time period part 1	1.021	1.010	1.021	1.013	1.021	1.012
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Time period part 2	0.964	1.036	0.964	1.047	0.964	1.031
	(0.001)	(0.032)	(0.001)	(0.032)	(0.001)	(0.032)
Time period part 3	1.010	1.001	1.010	1.002	1.010	1.001

	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)

Number of mortgages 6112

Table 37. Results comparison in Tampa

Prepayment and Default are Dependent Variables

		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and default		The credit score for default		The credit score for prepayment and default		The credit score for default	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	the negative equity dummy	0.9283	-21862.5	0.9283	-21869.4	0.9275	-21820	0.9274	-21824.9
2	original LTV, the negative equity dummy	0.9284	-21869	0.9283	-21876	0.9275	-21826.6	0.9275	-21831.5
3	negative equity, the negative equity dummy	0.9284	-21873.2	0.9283	-21880.2	0.9275	-21830.8	0.9274	-21835.7

Prepayment and 90-days-delinquency are Dependent Variables

		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and 90-days-delinquency		The credit score for 90-days-delinquency		The credit score for prepayment and 90-days-delinquency		The credit score for 90-days-delinquency	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	the negative equity dummy	0.9255	-22354.3	0.9254	-22358.3	0.9256	-22020.5	0.9256	-22022.6
2	original LTV, the negative equity dummy	0.9255	-22359.3	0.9255	-22363.2	0.9257	-22025.5	0.9256	-22027.6
3	negative equity	0.9255	-22366.6	0.9254	-22370.5	0.9256	-22031.9	0.9256	-22034

Table 38. Coefficients for Detroit: Prepayment and default are dependent variables, the credit score for both the prepayment and the default

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	original loan-to-value is one of explanatory variables		the negative equity dummy and original loan-to-value are explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	Default	Prepayment	Default	Prepayment	Default

Prepayment penalty	0.531	--	0.531	--	0.531	--
	(0.457)		(0.457)		(0.457)	
Call option variables						
Call option part 1	4.746	--	4.746	--	4.746	--
	(1.946)		(1.946)		(1.946)	
Call option part 2	22.409	--	22.409	--	22.409	--
	(1.871)		(1.871)		(1.871)	
Call option part 3	8.305	--	8.305	--	8.305	--
	(1.199)		(1.199)		(1.199)	
Call option part 4	0.776	--	0.776	--	0.776	--
	(0.850)		(0.850)		(0.850)	
The unemployment rate variables						
The unemployment rate part 1	--	0.209	--	0.213	--	0.217
		(0.172)		(0.172)		(0.172)
The unemployment rate part 2	--	-0.030	--	-0.036	--	-0.038
		(0.047)		(0.048)		(0.048)
Negative equity variables						
Negative equity part 1	--	--	--	--	--	0.048
						(0.028)
Negative equity part 2	--	--	--	--	--	-0.028
						(0.029)
Negative equity part 3	--	--	--	--	--	-0.002
						(0.024)
Negative equity part 4	--	--	--	--	--	-0.008
						(0.011)
The negative equity dummy	--	--	--	0.117	--	--
				(0.177)		
Original loan-to-value variables						
Original LTV part 1	--	5.569	--	5.414	--	5.214
		(1.669)		(1.676)		(1.661)
Original LTV part 2	--	3.607	--	3.242	--	2.772
		(2.806)		(2.858)		(2.886)
Original LTV part 3	--	0.426	--	0.286	--	1.127
		(7.196)		(7.204)		(7.267)
Original LTV part 4	--	5.311		5.227	--	5.479
		(1.190)		(1.196)		(1.235)
Original LTV part 5	--	2.423	--	2.090	--	3.198
		(7.104)		(7.119)		(7.211)
Log loan size variables						
Log loan size part 1	0.750	-1.095	0.750	-1.102	0.750	-1.101
	(0.186)	(0.272)	(0.186)	(0.272)	(0.186)	(0.278)
Log loan size part 2	1.014	-0.516	1.014	-0.528	1.014	-0.439
	(0.273)	(0.580)	(0.273)	(0.580)	(0.273)	(0.595)
Log loan size part 3	0.749	-1.666	0.749	-1.667	0.749	-1.528
	(0.248)	(0.682)	(0.248)	(0.682)	(0.248)	(0.697)
Log loan size part 4	1.324	-0.083	1.324	-0.081	1.324	0.096
	(0.155)	(0.546)	(0.155)	(0.546)	(0.155)	(0.583)
Loan age variables						
Loan age part 1	0.064	0.173	0.064	0.172	0.064	0.172
	(0.010)	(0.047)	(0.010)	(0.047)	(0.010)	(0.047)
Loan age part 2	-0.009	-0.012	-0.009	-0.013	-0.009	-0.013
	(0.005)	(0.014)	(0.005)	(0.015)	(0.005)	(0.015)
Loan age part 3	-0.003	0.011	-0.003	0.010	-0.003	0.010
	(0.004)	(0.008)	(0.004)	(0.008)	(0.004)	(0.008)
Loan age part 4	0.001	-0.000	0.001	0.000	0.001	-0.001
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
The credit score variables						
The credit score part 1	4.487	-7.495	4.487	-7.464	4.487	-7.305
	(1.486)	(2.296)	(1.486)	(2.296)	(1.486)	(2.309)
The credit score part 2	0.431	-7.211	0.431	-7.173	0.431	-7.162

	(1.793)	(3.934)	(1.793)	(3.933)	(1.793)	(3.938)
The credit score part 3	8.199	-9.115	8.199	-9.089	8.198	-9.614
	(2.232)	(6.581)	(2.232)	(6.579)	(2.232)	(6.594)
The credit score part 4	0.228	-20.403	0.228	-20.349	0.228	-19.982
	(3.016)	(12.295)	(3.016)	(12.299)	(3.016)	(12.315)
The debt-to-income ratio variables						
DTI part 1	-0.247	0.431	-0.247	0.447	-0.247	0.491
	(0.870)	(2.520)	(0.870)	(2.520)	(0.870)	(2.521)
DTI part 2	0.280	-0.456	0.280	-0.497	0.280	-0.729
	(1.139)	(3.159)	(1.139)	(3.159)	(1.139)	(3.164)
DTI part 3	-0.151	4.716	-0.151	4.737	-0.151	4.975
	(0.977)	(2.392)	(0.977)	(2.392)	(0.977)	(2.403)
DTI part 4	-0.653	2.140	-0.653	2.158	-0.653	2.214
	(0.698)	(1.259)	(0.698)	(1.259)	(0.698)	(1.255)
The dummy for the number of units	-0.421	0.652	-0.421	0.648	-0.421	0.665
	(0.229)	(0.333)	(0.229)	(0.334)	(0.229)	(0.334)
Time period variables						
Time period part 1	0.006	0.009	0.006	0.008	0.006	0.008
	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)	(0.005)
Time period part 2	-0.045	0.083	-0.045	0.082	-0.045	0.083
	(0.002)	(0.021)	(0.002)	(0.021)	(0.002)	(0.021)
Time period part 3	0.032	-0.045	0.032	-0.045	0.032	-0.045
	(0.003)	(0.012)	(0.003)	(0.012)	(0.003)	(0.012)

Number of mortgages 3167

Table 39. Odds ratio for Detroit: Prepayment and default are dependent variables, the credit score for both the prepayment and the default

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	original loan-to-value is one of explanatory variables		the negative equity dummy and original loan-to-value are explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	1.700	--	1.700	--	1.700	--
	(0.776)		(0.776)		(0.776)	
Call option variables						
Call option part 1	115.153	--	115.163	--	115.160	--
	(224.135)		(224.154)		(224.147)	
Call option part 2	5.396e+09	--	5.396e+09	--	5.396e+09	--
	(1.010e+10)		(1.010e+10)		(1.010e+10)	
Call option part 3	4,042.849	--	4,043.310	--	4,042.498	--
	(4,846.310)		(4,846.860)		(4,845.893)	
Call option part 4	2.173	--	2.172	--	2.172	--
	(1.847)		(1.846)		(1.846)	
The unemployment rate variables						
The unemployment rate part 1	--	1.232	--	1.237	--	1.242
		(0.212)		(0.213)		(0.213)
The unemployment rate part 2	--	0.970	--	0.965	--	0.963
		(0.046)		(0.046)		(0.046)
Negative equity variables						
Negative equity part 1	--	--	--	--	--	1.049
						(0.029)

Negative equity part 2	--	--	--	--	--	0.972
						(0.029)
Negative equity part 3	--	--	--	--	--	0.998
						(0.024)
Negative equity part 4	--	--	--	--	--	0.992
						(0.011)
The negative equity dummy	--	--	--	1.124	--	--
				(0.199)		
Original loan-to-value variables						
Original LTV part 1	--	262.101	--	224.434	--	183.774
		(437.505)		(376.177)		(305.247)
Original LTV part 2	--	36.864	--	25.595	--	15.997
		(103.424)		(73.146)		(46.168)
Original LTV part 3	--	1.531	--	1.331	--	3.086
		(11.018)		(9.587)		(22.426)
Original LTV part 4	--	202.522		186.289	--	239.644
		(240.970)		(222.832)		(295.902)
Original LTV part 5	--	11.278	--	8.083	--	24.488
		(80.119)		(57.539)		(176.591)
Log loan size variables						
Log loan size part 1	2.118	0.334	2.118	0.332	2.118	0.333
	(0.394)	(0.091)	(0.394)	(0.090)	(0.394)	(0.093)
Log loan size part 2	2.756	0.597	2.756	0.590	2.756	0.645
	(0.751)	(0.346)	(0.751)	(0.342)	(0.751)	(0.384)
Log loan size part 3	2.115	0.189	2.115	0.189	2.115	0.217
	(0.524)	(0.129)	(0.524)	(0.129)	(0.524)	(0.151)
Log loan size part 4	3.759	0.920	3.759	0.922	3.759	1.101
	(0.581)	(0.503)	(0.581)	(0.504)	(0.581)	(0.642)
Loan age variables						
Loan age part 1	1.066	1.189	1.066	1.188	1.066	1.188
	(0.010)	(0.056)	(0.010)	(0.056)	(0.010)	(0.056)
Loan age part 2	0.991	0.988	0.991	0.987	0.991	0.987
	(0.005)	(0.014)	(0.005)	(0.014)	(0.005)	(0.014)
Loan age part 3	0.997	1.011	0.997	1.010	0.997	1.010
	(0.004)	(0.008)	(0.004)	(0.008)	(0.004)	(0.008)
Loan age part 4	1.001	1.000	1.001	1.000	1.001	0.999
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
The credit score variables						
The credit score part 1	88.812	0.001	88.823	0.001	88.861	0.001
	(132.005)	(0.001)	(132.022)	(0.001)	(132.079)	(0.002)
The credit score part 2	1.539	0.001	1.539	0.001	1.539	0.001
	(2.759)	(0.003)	(2.759)	(0.003)	(2.759)	(0.003)
The credit score part 3	3,635.538	0.000	3,635.527	0.000	3,634.758	0.000
	(8,116.133)	(0.001)	(8,116.111)	(0.001)	(8,114.395)	(0.000)
The credit score part 4	1.256	0.000	1.256	0.000	1.256	0.000
	(3.787)	(0.000)	(3.787)	(0.000)	(3.789)	(0.000)
The debt-to-income ratio variables						
DTI part 1	0.781	1.539	0.781	1.563	0.781	1.634
	(0.679)	(3.879)	(0.679)	(3.938)	(0.679)	(4.120)
DTI part 2	1.323	0.634	1.323	0.608	1.323	0.483
	(1.506)	(2.001)	(1.506)	(1.921)	(1.506)	(1.527)
DTI part 3	0.860	111.698	0.860	114.045	0.860	144.728
	(0.840)	(267.210)	(0.840)	(272.851)	(0.840)	(347.819)
DTI part 4	0.520	8.498	0.520	8.650	0.521	9.154
	(0.363)	(10.701)	(0.363)	(10.889)	(0.363)	(11.489)
The dummy for the number of units	0.656	1.920	0.656	1.911	0.657	1.945
	(0.150)	(0.640)	(0.150)	(0.637)	(0.150)	(0.649)
Time period variables						
Time period part 1	1.006	1.009	1.006	1.009	1.006	1.008
	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)	(0.005)
Time period part 2	0.956	1.087	0.956	1.085	0.956	1.087
	(0.002)	(0.022)	(0.002)	(0.022)	(0.002)	(0.023)

Time period part 3	1.033 (0.003)	0.956 (0.012)	1.033 (0.003)	0.956 (0.012)	1.033 (0.003)	0.956 (0.012)

Number of mortgages 3167

Table 40. Coefficients for Detroit: Prepayment and 90-days-delinquency are dependent variables, the credit score for both the prepayment and the 90-days-delinquency

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	original loan-to-value is one of explanatory variables		the negative equity dummy and original loan-to-value are explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	0.524 (0.457)	--	0.524 (0.457)	--	0.524 (0.457)	--
Call option variables						
Call option part 1	4.520 (1.979)	--	4.520 (1.979)	--	4.520 (1.979)	--
Call option part 2	22.314 (1.920)	--	22.314 (1.920)	--	22.314 (1.920)	--
Call option part 3	9.101 (1.214)	--	9.101 (1.214)	--	9.100 (1.214)	--
Call option part 4	1.182 (0.851)	--	1.182 (0.851)	--	1.182 (0.851)	--
The unemployment rate variables						
The unemployment rate part 1	--	0.088 (0.131)	--	0.090 (0.131)	--	0.090 (0.131)
The unemployment rate part 2	--	0.058 (0.042)	--	0.055 (0.043)	--	0.061 (0.043)
Negative equity variables						
Negative equity part 1	--	--	--	--	--	0.020 (0.027)
Negative equity part 2	--	--	--	--	--	-0.023 (0.029)
Negative equity part 3	--	--	--	--	--	-0.004 (0.023)
Negative equity part 4	--	--	--	--	--	-0.005 (0.010)
The negative equity dummy	--	--	--	0.055 (0.165)	--	--
Original loan-to-value variables						
Original LTV part 1	--	5.318 (1.571)	--	5.258 (1.578)	--	5.169 (1.574)
Original LTV part 2	--	2.635 (2.373)	--	2.469 (2.425)	--	2.471 (2.446)
Original LTV part 3	--	2.800 (6.591)	--	2.758 (6.595)	--	3.730 (6.661)
Original LTV part 4	--	5.377 (1.090)	--	5.337 (1.096)	--	5.612 (1.125)
Original LTV part 5	--	1.115 (6.520)	--	0.936 (6.540)	--	2.176 (6.641)
Log loan size variables						

Log loan size part 1	0.799	-0.813	0.799	-0.816	0.799	-0.797
	(0.190)	(0.263)	(0.190)	(0.263)	(0.190)	(0.268)
Log loan size part 2	1.006	-0.685	1.006	-0.689	1.006	-0.598
	(0.276)	(0.536)	(0.276)	(0.536)	(0.276)	(0.547)
Log loan size part 3	0.780	-1.301	0.780	-1.303	0.780	-1.191
	(0.249)	(0.620)	(0.249)	(0.620)	(0.249)	(0.630)
Log loan size part 4	1.308	-0.330	1.308	-0.329	1.308	-0.207
	(0.155)	(0.506)	(0.155)	(0.506)	(0.155)	(0.541)
Loan age variables						
Loan age part 1	0.064	0.130	0.064	0.130	0.064	0.130
	(0.010)	(0.035)	(0.010)	(0.035)	(0.010)	(0.035)
Loan age part 2	-0.008	-0.003	-0.008	-0.003	-0.008	-0.002
	(0.005)	(0.013)	(0.005)	(0.013)	(0.005)	(0.013)
Loan age part 3	-0.004	0.003	-0.004	0.002	-0.004	0.004
	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.007)
Loan age part 4	0.002	0.003	0.002	0.003	0.002	0.002
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
The credit score variables						
The credit score part 1	4.491	-8.282	4.491	-8.276	4.491	-8.164
	(1.468)	(2.026)	(1.468)	(2.026)	(1.468)	(2.036)
The credit score part 2	-0.295	-13.044	-0.295	-13.014	-0.295	-13.129
	(1.831)	(3.843)	(1.831)	(3.843)	(1.831)	(3.850)
The credit score part 3	8.540	-3.255	8.540	-3.255	8.540	-3.473
	(2.247)	(6.303)	(2.247)	(6.302)	(2.247)	(6.314)
The credit score part 4	-0.172	-27.148	-0.172	-27.105	-0.172	-27.043
	(3.115)	(12.545)	(3.115)	(12.546)	(3.115)	(12.567)
The debt-to-income ratio variables						
DTI part 1	-0.124	0.787	-0.124	0.793	-0.124	0.780
	(0.872)	(2.371)	(0.872)	(2.371)	(0.872)	(2.372)
DTI part 2	0.227	1.160	0.227	1.143	0.227	1.040
	(1.142)	(2.862)	(1.142)	(2.862)	(1.142)	(2.865)
DTI part 3	-0.197	3.890	-0.197	3.892	-0.197	4.063
	(0.985)	(2.152)	(0.985)	(2.152)	(0.985)	(2.160)
DTI part 4	-0.535	1.677	-0.535	1.687	-0.535	1.712
	(0.702)	(1.199)	(0.702)	(1.199)	(0.702)	(1.198)
The dummy for the number of units	-0.479	0.659	-0.479	0.658	-0.479	0.662
	(0.235)	(0.304)	(0.235)	(0.304)	(0.235)	(0.305)
Time period variables						
Time period part 1	0.007	0.008	0.007	0.008	0.007	0.007
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Time period part 2	-0.045	0.040	-0.045	0.039	-0.045	0.041
	(0.002)	(0.019)	(0.002)	(0.019)	(0.002)	(0.019)
Time period part 3	0.031	-0.035	0.031	-0.035	0.031	-0.035
	(0.003)	(0.012)	(0.003)	(0.012)	(0.003)	(0.012)

Number of mortgages 3152

Table 41. Odds ratios for Detroit: Prepayment and 90-days-delinquency are dependent variables, the credit score for both the prepayment and the 90-days-delinquency

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	original loan-to-value is one of explanatory variables		the negative equity dummy and original loan-to-value are explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency

Prepayment penalty	1.689 (0.771)	--	1.689 (0.771)	--	1.689 (0.771)	--
Call option variables						
Call option part 1	91.851 (181.808)	--	91.855 (181.817)	--	91.846 (181.799)	--
Call option part 2	4.907e+09 (9.420e+09)	--	4.907e+09 (9.420e+09)	--	4.907e+09 (9.419e+09)	--
Call option part 3	8,960.957 (10,880.620)	--	8,961.448 (10,881.210)	--	8,958.873 (10,878.110)	--
Call option part 4	3.260 (2.775)	--	3.260 (2.775)	--	3.261 (2.775)	--
The unemployment rate variables						
The unemployment rate part 1	--	1.092 (0.143)	--	1.095 (0.143)	--	1.095 (0.143)
The unemployment rate part 2	--	1.059 (0.044)	--	1.056 (0.045)	--	1.062 (0.046)
Negative equity variables						
Negative equity part 1	--	--	--	--	--	1.021 (0.027)
Negative equity part 2	--	--	--	--	--	0.977 (0.028)
Negative equity part 3	--	--	--	--	--	0.996 (0.023)
Negative equity part 4	--	--	--	--	--	0.995 (0.010)
The negative equity dummy	--	--	--	1.057 (0.174)	--	--
Original loan-to-value variables						
Original LTV part 1	--	204.052 (320.611)	--	192.021 (303.011)	--	175.697 (276.499)
Original LTV part 2	--	13.949 (33.104)	--	11.809 (28.633)	--	11.836 (28.955)
Original LTV part 3	--	16.450 (108.424)	--	15.762 (103.951)	--	41.696 (277.757)
Original LTV part 4	--	216.345 (235.805)	--	207.952 (227.934)	--	273.779 (307.935)
Original LTV part 5	--	3.049 (19.881)	--	2.550 (16.678)	--	8.815 (58.538)
Log loan size variables						
Log loan size part 1	2.223 (0.422)	0.443 (0.117)	2.223 (0.422)	0.442 (0.116)	2.223 (0.422)	0.451 (0.121)
Log loan size part 2	2.735 (0.754)	0.504 (0.270)	2.735 (0.754)	0.502 (0.269)	2.735 (0.754)	0.550 (0.301)
Log loan size part 3	2.182 (0.543)	0.272 (0.169)	2.182 (0.543)	0.272 (0.168)	2.182 (0.543)	0.304 (0.192)
Log loan size part 4	3.698 (0.572)	0.719 (0.364)	3.698 (0.572)	0.720 (0.364)	3.698 (0.572)	0.813 (0.440)
Loan age variables						
Loan age part 1	1.066 (0.010)	1.139 (0.040)	1.066 (0.010)	1.139 (0.040)	1.066 (0.010)	1.139 (0.040)
Loan age part 2	0.992 (0.005)	0.997 (0.013)	0.992 (0.005)	0.997 (0.013)	0.992 (0.005)	0.998 (0.013)
Loan age part 3	0.996 (0.004)	1.003 (0.007)	0.996 (0.004)	1.002 (0.007)	0.996 (0.004)	1.004 (0.007)
Loan age part 4	1.002 (0.002)	1.003 (0.004)	1.002 (0.002)	1.003 (0.004)	1.002 (0.002)	1.002 (0.004)
The credit score variables						
The credit score part 1	89.172	0.000	89.179	0.000	89.195	0.000

	(130.868)	(0.001)	(130.878)	(0.001)	(130.901)	(0.001)
The credit score part 2	0.744	0.000	0.744	0.000	0.744	0.000
	(1.363)	(0.000)	(1.363)	(0.000)	(1.363)	(0.000)
The credit score part 3	5,114.252	0.039	5,114.199	0.039	5,113.635	0.031
	(11,490.884)	(0.243)	(11,490.768)	(0.243)	(11,489.499)	(0.196)
The credit score part 4	0.842	0.000	0.842	0.000	0.842	0.000
	(2.621)	(0.000)	(2.621)	(0.000)	(2.623)	(0.000)
The debt-to-income ratio variables						
DTI part 1	0.883	2.197	0.883	2.211	0.883	2.182
	(0.770)	(5.207)	(0.770)	(5.241)	(0.770)	(5.176)
DTI part 2	1.255	3.191	1.255	3.137	1.255	2.828
	(1.433)	(9.134)	(1.433)	(8.979)	(1.433)	(8.102)
DTI part 3	0.821	48.924	0.821	49.016	0.821	58.156
	(0.809)	(105.275)	(0.809)	(105.469)	(0.809)	(125.615)
DTI part 4	0.586	5.348	0.586	5.405	0.586	5.540
	(0.411)	(6.412)	(0.411)	(6.481)	(0.411)	(6.639)
The dummy for the number of units	0.620	1.933	0.620	1.930	0.620	1.938
	(0.146)	(0.588)	(0.146)	(0.588)	(0.146)	(0.591)
Time period variables						
Time period part 1	1.007	1.008	1.007	1.008	1.007	1.007
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Time period part 2	0.956	1.040	0.956	1.040	0.956	1.042
	(0.002)	(0.020)	(0.002)	(0.020)	(0.002)	(0.020)
Time period part 3	1.032	0.966	1.032	0.966	1.032	0.965
	(0.003)	(0.012)	(0.003)	(0.012)	(0.003)	(0.012)

Number of mortgages 3152

Table 42. Results comparison in Detroit

Prepayment and Default are Dependent Variables

		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and default		The credit score for default		The credit score for prepayment and default		The credit score for default	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	original LTV	0.9246	-11836.532	0.9244	-11858.5	0.9236	-11798.998	0.9233	-11818.9
2	original LTV, the negative equity dummy	0.9246	-11840.344	0.9244	-11862.3	0.9236	-11802.819	0.9233	-11822.7
3	negative equity, original LTV	0.9247	-11850.587	0.9244	-11872.5	0.9236	-11813.206	0.9233	-11833.1

Prepayment and 90-days-delinquency are Dependent Variables

Right censored dataset		Deleting modification dataset	
The credit score for prepayment and 90-days-delinquency	The credit score for 90-days-delinquency	The credit score for prepayment and 90-days-delinquency	The credit score for 90-days-delinquency

Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	original LTV	0.9211	-12034.2	0.9209	-12052.9	0.9212	-11841.5	0.921	-11858.5
2	original LTV, the negative equity dummy	0.9211	-12038.2	0.9209	-12056.9	0.9212	-11845.2	0.921	-11862.2
3	negative equity, original LTV	0.9211	-12049.4	0.9209	-12068.1	0.9212	-11855.2	0.921	-11872.2

Table 43. Coefficients for Las Vegas: Prepayment and default are dependent variables, the credit score only for default

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	negative equity is one of explanatory variables		negative equity and the negative equity dummy are explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.674 (0.504)	--	-0.674 (0.504)	--	-0.675 (0.504)	--
Call option variables						
Call option part 1	2.166 (1.404)	--	2.166 (1.404)	--	2.166 (1.404)	--
Call option part 2	4.922 (1.539)	--	4.923 (1.539)	--	4.923 (1.539)	--
Call option part 3	9.522 (1.232)	--	9.522 (1.232)	--	9.521 (1.232)	--
Call option part 4	2.995 (0.766)	--	2.995 (0.766)	--	2.996 (0.766)	--
The unemployment rate variables						
The unemployment rate part 1	--	0.653 (0.227)	--	0.637 (0.227)	--	0.652 (0.227)
The unemployment rate part 2	--	0.014 (0.098)	--	0.012 (0.098)	--	0.028 (0.098)
Negative equity variables						
Negative equity part 1	--	0.035 (0.011)	--	0.015 (0.020)	--	0.032 (0.011)
Negative equity part 2	--	0.025 (0.011)	--	0.028 (0.011)	--	0.024 (0.011)
Negative equity part 3	--	0.022 (0.006)	--	0.022 (0.006)	--	0.020 (0.006)
Negative equity part 4	--	0.006 (0.003)	--	0.006 (0.003)	--	0.003 (0.004)
The negative equity dummy	--	--	--	0.404 (0.360)	--	--
Original loan-to-value variables						
Original LTV part 1	--	--	--	--	--	1.080 (1.419)
Original LTV part 2	--	--	--	--	--	-1.350 (1.792)
Original LTV part 3	--	--	--	--	--	10.099 (8.380)

Original LTV part 4	--	--	--	--	--	1.625
						(1.020)
Original LTV part 5	--	--	--	--	--	1.140
						(7.401)
Log loan size variables						
Log loan size part 1	0.657	0.372	0.657	0.353	0.657	0.247
	(0.138)	(0.580)	(0.138)	(0.582)	(0.138)	(0.589)
Log loan size part 2	0.948	1.146	0.948	1.167	0.948	1.178
	(0.224)	(0.704)	(0.224)	(0.704)	(0.224)	(0.705)
Log loan size part 3	-0.101	-0.637	-0.101	-0.643	-0.101	-0.560
	(0.249)	(0.563)	(0.249)	(0.563)	(0.249)	(0.563)
Log loan size part 4	0.404	-1.217	0.404	-1.215	0.404	-0.989
	(0.220)	(0.431)	(0.220)	(0.431)	(0.220)	(0.439)
Loan age variables						
Loan age part 1	0.166	0.605	0.166	0.605	0.166	0.599
	(0.015)	(0.291)	(0.015)	(0.292)	(0.015)	(0.291)
Loan age part 2	0.001	0.078	0.001	0.077	0.001	0.080
	(0.006)	(0.024)	(0.006)	(0.024)	(0.006)	(0.024)
Loan age part 3	0.001	-0.018	0.001	-0.018	0.001	-0.013
	(0.004)	(0.008)	(0.004)	(0.008)	(0.004)	(0.008)
Loan age part 4	0.005	0.001	0.005	0.002	0.005	0.001
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
The credit score variables						
The credit score part 1	--	-4.319	--	-4.219	--	-4.664
		(2.618)		(2.619)		(2.643)
The credit score part 2	--	-0.282	--	-0.334	--	-0.466
		(3.584)		(3.585)		(3.592)
The credit score part 3	--	-11.989	--	-11.938	--	-11.518
		(5.191)		(5.189)		(5.187)
The credit score part 4	--	-17.205	--	-17.193	--	-17.356
		(8.697)		(8.696)		(8.712)
The debt-to-income ratio variables						
DTI part 1	0.739	0.671	0.739	0.655	0.739	0.838
	(0.582)	(1.782)	(0.582)	(1.780)	(0.582)	(1.788)
DTI part 2	-1.280	2.275	-1.280	2.312	-1.279	2.279
	(0.871)	(2.175)	(0.871)	(2.174)	(0.871)	(2.176)
DTI part 3	-0.720	1.049	-0.720	1.025	-0.720	0.849
	(0.882)	(1.856)	(0.882)	(1.856)	(0.882)	(1.859)
DTI part 4	0.265	-0.400	0.265	-0.432	0.265	-0.463
	(0.723)	(1.263)	(0.723)	(1.264)	(0.723)	(1.261)
The dummy for the number of units	-0.575	-0.320	-0.575	-0.322	-0.575	-0.278
	(0.309)	(0.713)	(0.309)	(0.713)	(0.309)	(0.715)
Time period variables						
Time period part 1	0.035	-0.009	0.035	-0.009	0.035	-0.007
	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)	(0.005)
Time period part 2	-0.044	0.055	-0.044	0.056	-0.044	0.051
	(0.001)	(0.030)	(0.001)	(0.030)	(0.001)	(0.030)
Time period part 3	0.016	-0.030	0.016	-0.030	0.016	-0.028
	(0.002)	(0.015)	(0.002)	(0.015)	(0.002)	(0.015)
Time period part 4	--	0.047	--	0.047	--	0.048
		(0.027)		(0.027)		(0.027)

Number of mortgages 4487

Table 44. Odds ratio for Las Vegas: Prepayment and default are dependent variables, the credit score only for default

	#1 dataset with right censored mortgages	#1 dataset with right censored mortgages	#1 dataset with right censored mortgages
	negative equity is one of explanatory variables	negative equity and the negative equity dummy are explanatory	negative equity and original loan-to-value are explanatory

			variables		variables	
	Prepayment	Default	Prepayment	Default	Prepayment	Default
Prepayment penalty	0.509	--	0.509	--	0.509	--
	(0.257)		(0.257)		(0.257)	
Call option variables						
Call option part 1	8.725	--	8.725	--	8.725	--
	(12.249)		(12.248)		(12.248)	
Call option part 2	137.345	--	137.352	--	137.436	--
	(211.397)		(211.407)		(211.537)	
Call option part 3	13,661.278	--	13,659.570	--	13,648.683	--
	(16,835.353)		(16,833.252)		(16,819.863)	
Call option part 4	19.976	--	19.982	--	20.005	--
	(15.309)		(15.313)		(15.332)	
The unemployment rate variables						
The unemployment rate part 1	--	1.922	--	1.890	--	1.919
		(0.435)		(0.430)		(0.435)
The unemployment rate part 2	--	1.014	--	1.013	--	1.029
		(0.100)		(0.099)		(0.101)
Negative equity variables						
Negative equity part 1	--	1.035	--	1.016	--	1.033
		(0.012)		(0.021)		(0.012)
Negative equity part 2	--	1.025	--	1.028	--	1.024
		(0.011)		(0.012)		(0.011)
Negative equity part 3	--	1.022	--	1.022	--	1.021
		(0.006)		(0.006)		(0.007)
Negative equity part 4	--	1.006	--	1.006	--	1.003
		(0.003)		(0.003)		(0.004)
The negative equity dummy	--	--	--	1.497	--	--
				(0.539)		
Original loan-to-value variables						
Original LTV part 1	--	--	--	--	--	2.944
						(4.177)
Original LTV part 2	--	--	--	--	--	0.259
						(0.465)
Original LTV part 3	--	--	--	--	--	24,307.274
						(203,693.947)
Original LTV part 4	--	--	--	--	--	5.079
						(5.181)
Original LTV part 5	--	--	--	--	--	3.125
						(23.131)
Log loan size variables						
Log loan size part 1	1.930	1.450	1.930	1.423	1.930	1.281
	(0.266)	(0.842)	(0.266)	(0.829)	(0.266)	(0.755)
Log loan size part 2	2.580	3.145	2.580	3.212	2.580	3.247
	(0.578)	(2.214)	(0.578)	(2.261)	(0.578)	(2.289)
Log loan size part 3	0.904	0.529	0.904	0.526	0.904	0.571
	(0.225)	(0.298)	(0.225)	(0.296)	(0.225)	(0.322)
Log loan size part 4	1.497	0.296	1.497	0.297	1.497	0.372
	(0.329)	(0.127)	(0.329)	(0.128)	(0.329)	(0.163)
Loan age variables						
Loan age part 1	1.180	1.832	1.180	1.831	1.180	1.821
	(0.017)	(0.534)	(0.017)	(0.534)	(0.017)	(0.529)
Loan age part 2	1.001	1.081	1.001	1.080	1.001	1.084
	(0.006)	(0.026)	(0.006)	(0.026)	(0.006)	(0.026)
Loan age part 3	1.001	0.982	1.001	0.982	1.001	0.987
	(0.004)	(0.008)	(0.004)	(0.008)	(0.004)	(0.008)
Loan age part 4	1.005	1.001	1.005	1.002	1.005	1.001

	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
The credit score variables						
The credit score part 1	--	0.013	--	0.015	--	0.009
		(0.035)		(0.039)		(0.025)
The credit score part 2	--	0.755	--	0.716	--	0.627
		(2.704)		(2.568)		(2.253)
The credit score part 3	--	0.000	--	0.000	--	0.000
		(0.000)		(0.000)		(0.000)
The credit score part 4	--	0.000	--	0.000	--	0.000
		(0.000)		(0.000)		(0.000)
The debt-to-income ratio variables						
DTI part 1	2.094	1.955	2.094	1.924	2.094	2.312
	(1.219)	(3.485)	(1.219)	(3.425)	(1.219)	(4.133)
DTI part 2	0.278	9.725	0.278	10.096	0.278	9.767
	(0.242)	(21.149)	(0.242)	(21.949)	(0.242)	(21.251)
DTI part 3	0.487	2.853	0.487	2.787	0.487	2.338
	(0.429)	(5.295)	(0.429)	(5.171)	(0.429)	(4.346)
DTI part 4	1.303	0.670	1.303	0.649	1.304	0.630
	(0.943)	(0.847)	(0.943)	(0.820)	(0.943)	(0.794)
The dummy for the number of units	0.563	0.726	0.563	0.725	0.563	0.757
	(0.174)	(0.518)	(0.174)	(0.517)	(0.174)	(0.541)
Time period variables						
Time period part 1	1.036	0.991	1.036	0.991	1.036	0.993
	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)	(0.005)
Time period part 2	0.957	1.057	0.957	1.057	0.957	1.053
	(0.001)	(0.031)	(0.001)	(0.031)	(0.001)	(0.031)
Time period part 3	1.016	0.971	1.016	0.971	1.016	0.972
	(0.002)	(0.014)	(0.002)	(0.014)	(0.002)	(0.014)
Time period part 4	--	1.048	--	1.048	--	1.050
		(0.029)		(0.029)		(0.029)

Number of mortgages 4487

Table 45. Coefficients for Las Vegas: Prepayment and 90-days-delinquency are dependent variables, the credit score only for 90-days-delinquency

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	negative equity is one of explanatory variables		negative equity and the negative equity dummy are explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	-0.676	--	-0.676	--	-0.676	--
	(0.504)		(0.504)		(0.504)	
Call option variables						
Call option part 1	2.120	--	2.120	--	2.120	--
	(1.426)		(1.426)		(1.426)	
Call option part 2	4.970	--	4.970	--	4.970	--
	(1.565)		(1.565)		(1.565)	
Call option part 3	9.866	--	9.866	--	9.866	--
	(1.240)		(1.240)		(1.240)	
Call option part 4	3.302	--	3.302	--	3.304	--
	(0.763)		(0.763)		(0.763)	
The unemployment rate variables						

The unemployment rate part 1	--	0.603	--	0.557	--	0.589
		(0.178)		(0.180)		(0.178)
The unemployment rate part 2	--	0.045	--	0.041	--	0.061
		(0.089)		(0.089)		(0.090)
Negative equity variables						
Negative equity part 1	--	0.049	--	0.003	--	0.047
		(0.011)		(0.018)		(0.011)
Negative equity part 2	--	0.024	--	0.032	--	0.023
		(0.010)		(0.011)		(0.010)
Negative equity part 3	--	0.020	--	0.019	--	0.018
		(0.006)		(0.006)		(0.006)
Negative equity part 4	--	0.005	--	0.005	--	0.002
		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	--	--	0.977	--	
				(0.316)		
Original loan-to-value variables						
Original LTV part 1	--	--	--	--	--	0.444
						(1.295)
Original LTV part 2	--	--	--	--	--	-0.976
						(1.663)
Original LTV part 3	--	--	--	--	--	7.206
						(7.833)
Original LTV part 4	--	--	--	--	--	1.504
						(0.970)
Original LTV part 5	--	--	--	--	--	5.300
						(6.384)
Log loan size variables						
Log loan size part 1	0.661	0.119	0.661	0.062	0.661	0.031
	(0.140)	(0.525)	(0.140)	(0.529)	(0.140)	(0.535)
Log loan size part 2	1.041	1.261	1.041	1.316	1.041	1.304
	(0.219)	(0.652)	(0.219)	(0.652)	(0.219)	(0.653)
Log loan size part 3	-0.168	-0.676	-0.168	-0.690	-0.169	-0.607
	(0.252)	(0.538)	(0.252)	(0.538)	(0.252)	(0.538)
Log loan size part 4	0.423	-0.746	0.423	-0.741	0.423	-0.551
	(0.218)	(0.380)	(0.218)	(0.381)	(0.218)	(0.388)
Loan age variables						
Loan age part 1	0.164	0.350	0.164	0.348	0.164	0.345
	(0.015)	(0.101)	(0.015)	(0.101)	(0.015)	(0.101)
Loan age part 2	0.002	0.029	0.002	0.029	0.002	0.033
	(0.006)	(0.018)	(0.006)	(0.018)	(0.006)	(0.018)
Loan age part 3	0.001	-0.024	0.001	-0.025	0.001	-0.019
	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.008)
Loan age part 4	0.005	0.003	0.005	0.003	0.005	0.002
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
The credit score variables						
The credit score part 1	--	-5.413	--	-5.184	--	-5.648
		(2.395)		(2.395)		(2.421)
The credit score part 2	--	-0.943	--	-1.089	--	-0.938
		(3.371)		(3.371)		(3.380)
The credit score part 3	--	-12.784	--	-12.618	--	-12.417
		(4.844)		(4.840)		(4.843)
The credit score part 4	--	-13.987	--	-13.964	--	-14.219
		(8.413)		(8.407)		(8.434)
The debt-to-income ratio variables						
DTI part 1	0.896	0.373	0.895	0.331	0.896	0.460
	(0.585)	(1.680)	(0.585)	(1.674)	(0.585)	(1.685)
DTI part 2	-1.488	3.688	-1.488	3.787	-1.488	3.734
	(0.875)	(2.050)	(0.875)	(2.048)	(0.875)	(2.051)
DTI part 3	-0.683	-0.142	-0.683	-0.198	-0.683	-0.411
	(0.885)	(1.741)	(0.885)	(1.740)	(0.885)	(1.745)
DTI part 4	0.340	-0.545	0.340	-0.624	0.340	-0.594

	(0.722)	(1.207)	(0.722)	(1.207)	(0.722)	(1.203)
The dummy for the number of units	-0.601	-0.398	-0.601	-0.399	-0.601	-0.374
	(0.309)	(0.713)	(0.309)	(0.713)	(0.309)	(0.714)
Time period variables						
Time period part 1	0.035	-0.007	0.035	-0.008	0.035	-0.006
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Time period part 2	-0.043	0.030	-0.043	0.031	-0.043	0.026
	(0.001)	(0.028)	(0.001)	(0.028)	(0.001)	(0.028)
Time period part 3	0.015	-0.034	0.015	-0.033	0.015	-0.032
	(0.002)	(0.014)	(0.002)	(0.014)	(0.002)	(0.014)
Time period part 4	--	0.039	--	0.039	--	0.041
		(0.027)		(0.027)		(0.027)

Number of mortgages 4473

Table 46. Odds ratios for Las Vegas: Prepayment and 90-days-delinquency are dependent variables, the credit score only for 90-days-delinquency

	#1 dataset with right censored mortgages		#1 dataset with right censored mortgages		#1 dataset with right censored mortgages	
	negative equity is one of explanatory variables		negative equity and the negative equity dummy are explanatory variables		negative equity and original loan-to-value are explanatory variables	
	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency	Prepayment	90-days-delinquency
Prepayment penalty	0.509	--	0.509	--	0.508	--
	(0.256)		(0.256)		(0.256)	
Call option variables						
Call option part 1	8.333	--	8.332	--	8.333	--
	(11.879)		(11.877)		(11.878)	
Call option part 2	143.972	--	143.987	--	144.064	--
	(225.292)		(225.317)		(225.437)	
Call option part 3	19,270.157	--	19,264.699	--	19,257.001	--
	(23,894.858)		(23,888.093)		(23,878.565)	
Call option part 4	27.163	--	27.180	--	27.210	--
	(20.715)		(20.728)		(20.751)	
The unemployment rate variables						
The unemployment rate part 1	--	1.828	--	1.745	--	1.803
		(0.326)		(0.314)		(0.322)
The unemployment rate part 2	--	1.046	--	1.042	--	1.062
		(0.093)		(0.093)		(0.095)
Negative equity variables						
Negative equity part 1	--	1.050	--	1.003	--	1.048
		(0.011)		(0.018)		(0.011)
Negative equity part 2	--	1.024	--	1.032	--	1.023
		(0.011)		(0.011)		(0.011)
Negative equity part 3	--	1.020	--	1.019	--	1.019
		(0.006)		(0.006)		(0.006)
Negative equity part 4	--	1.005	--	1.005	--	1.002
		(0.003)		(0.003)		(0.003)
The negative equity dummy	--	--	--	2.656	--	
				(0.838)		
Original loan-to-value variables						
Original LTV part 1	--	--	--	--	--	1.558

						(2.018)
Original LTV part 2	--	--	--	--	--	0.377
						(0.627)
Original LTV part 3	--	--	--	--	--	1,347.897
						(10,557.942)
Original LTV part 4	--	--	--	--	--	4.501
						(4.368)
Original LTV part 5	--	--	--	--	--	200.299
						(1,278.672)
Log loan size variables						
Log loan size part 1	1.938	1.126	1.938	1.064	1.938	1.032
	(0.271)	(0.591)	(0.271)	(0.563)	(0.271)	(0.552)
Log loan size part 2	2.833	3.530	2.833	3.727	2.833	3.686
	(0.621)	(2.302)	(0.622)	(2.430)	(0.622)	(2.405)
Log loan size part 3	0.845	0.508	0.845	0.502	0.845	0.545
	(0.213)	(0.273)	(0.213)	(0.270)	(0.213)	(0.293)
Log loan size part 4	1.527	0.474	1.527	0.477	1.527	0.577
	(0.333)	(0.180)	(0.333)	(0.181)	(0.333)	(0.224)
Loan age variables						
Loan age part 1	1.178	1.419	1.178	1.416	1.178	1.412
	(0.017)	(0.143)	(0.017)	(0.144)	(0.017)	(0.143)
Loan age part 2	1.002	1.030	1.002	1.029	1.002	1.033
	(0.006)	(0.018)	(0.006)	(0.018)	(0.006)	(0.018)
Loan age part 3	1.001	0.977	1.001	0.976	1.001	0.981
	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.008)
Loan age part 4	1.005	1.003	1.005	1.003	1.005	1.002
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
The credit score variables						
The credit score part 1	--	0.004	--	0.006	--	0.004
		(0.011)		(0.013)		(0.009)
The credit score part 2	--	0.389	--	0.337	--	0.392
		(1.312)		(1.135)		(1.324)
The credit score part 3	--	0.000	--	0.000	--	0.000
		(0.000)		(0.000)		(0.000)
The credit score part 4	--	0.000	--	0.000	--	0.000
		(0.000)		(0.000)		(0.000)
The debt-to-income ratio variables						
DTI part 1	2.449	1.451	2.449	1.392	2.449	1.585
	(1.432)	(2.438)	(1.432)	(2.331)	(1.432)	(2.671)
DTI part 2	0.226	39.981	0.226	44.140	0.226	41.864
	(0.197)	(81.945)	(0.197)	(90.397)	(0.198)	(85.874)
DTI part 3	0.505	0.868	0.505	0.821	0.505	0.663
	(0.447)	(1.511)	(0.447)	(1.428)	(0.447)	(1.157)
DTI part 4	1.405	0.580	1.405	0.536	1.406	0.552
	(1.015)	(0.700)	(1.015)	(0.647)	(1.015)	(0.664)
The dummy for the number of units	0.548	0.672	0.548	0.671	0.548	0.688
	(0.170)	(0.479)	(0.170)	(0.478)	(0.170)	(0.491)
Time period variables						
Time period part 1	1.035	0.994	1.035	0.992	1.035	0.994
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Time period part 2	0.958	1.030	0.958	1.031	0.958	1.026
	(0.001)	(0.028)	(0.001)	(0.029)	(0.001)	(0.028)
Time period part 3	1.015	0.967	1.015	0.967	1.015	0.969
	(0.002)	(0.014)	(0.002)	(0.014)	(0.002)	(0.014)
Time period part 4	--	1.040	--	1.040	--	1.042
		(0.028)		(0.028)		(0.028)

Number of mortgages 4473

Table 47. Results comparison in Las Vegas

Prepayment and Default are Dependent Variables

		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and default		The credit score for default		The credit score for prepayment and default		The credit score for default	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	negative equity	0.9243	-15460.4	0.9243	-15453.3	0.9232	-15416.7	0.9231	-15408.7
2	negative equity and the negative equity dummy	0.9243	-15464	0.9243	-15456.9	0.9232	-15420.3	0.9231	-15412.3
3	negative equity and original LTV	0.9243	-15477.9	0.9243	-15470.8	0.9232	-15434.2	0.9231	-15426.1

Prepayment and 90-days-delinquency are Dependent Variables

		Right censored dataset				Deleting modification dataset			
		The credit score for prepayment and 90-days-delinquency		The credit score for 90-days-delinquency		The credit score for prepayment and 90-days-delinquency		The credit score for 90-days-delinquency	
Rank	Certain explanatory variables	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC	Pseudo R2	BIC
1	negative equity	0.9218	-15691	0.9218	-15681.7	0.9216	-15476.2	0.9215	-15466.1
2	negative equity and the negative equity dummy	0.9218	-15691	0.9218	-15681.7	0.9216	-15476.8	0.9216	-15466.7
3	negative equity and original LTV	0.9218	-15708.7	0.9218	-15699.4	0.9216	-15494.5	0.9216	-15484.4

equity. In Detroit, the value of negative equity and the negative equity dummy have an insignificant effect on the default decision, in conflict with hypotheses H4 and H5. Therefore, for the default/90-days-delinquency decision, negative equity is not a determinant factor.

In Miami and Detroit, the consistently positive effect of the loan-to-value supports H6 that mortgages with a higher LTV (a lower down-payment) are more likely to default (have delinquency). Moreover, in Detroit, the effect of the original loan-to-value are more significant compared with the effect of the value of negative equity, which makes it become a more determinant factor for default/delinquent decision. In Tampa and Las Vegas, the results of the original loan-to-value partially support H6. In Tampa, when the original loan-to-value is above the third quartile, its effect becomes insignificantly negative. In Las, Vegas, when the original loan-to-value is above the first quartile and is below the second quartile, its effect becomes insignificantly negative.

In Miami and Las Vegas, the credit score is only used to explain the default/90-days-delinquency risk. The negative effect of the credit score in these two MSAs supports H7 that households with higher credit score are less likely to default their mortgages. In Tampa and Detroit, the credit score is used to explain both termination risks. The effect of the credit score on the default/90-days-delinquency risk is negative which supports H7 and its effect on the prepayment risk is positive which indicates that households with higher credit score are more likely to prepay their mortgages.

In Miami, Tampa, and Las Vegas, the debt-to-income ratio has no clear effect on both the prepayment risk and the default/90-days-delinquency risk. In Detroit, the debt-to-income ratio has a positive effect on the default/90-days-delinquency risk and has no clear effect on the prepayment risk.

To argue that the results of Miami, Tampa, Detroit and Las Vegas are significantly different from those for Phoenix, the results of these four MSAs are compared with the 95% confident interval of the coefficients for Phoenix. The models used for comparison include the value of negative equity, the negative equity dummy and the original loan-to-value as explanatory variables. Moreover, the credit score is used to explain both prepayment and default/90-days-delinquency risk. The comparison results are listed in table 48.

[insert Table 48 here]

The main comparison results show that, for Miami, the coefficients of the second and the fourth spline of the value of call option, the coefficient of the first spline of the unemployment rate and the coefficient of the second spline of the value of negative equity for the 90-days-delinquency risk, the coefficients of the third and the fourth spline of original loan-to-value, the coefficients of log loan size for prepayment risk, the coefficients of loan age for the default and 90-days-delinquency risk, the coefficients of the first and the third spline of the credit score, and the coefficients of the debt-to-income ratio for the default and 90-days-delinquency risk are outside the 95% confidence interval of the results in Phoenix.

For Tampa, the coefficient of the first spline of the value of call option, the coefficient of the first spline of the unemployment rate for default, the coefficient of the second spline of the unemployment rate for 90-days-delinquency, the coefficient of the third spline of the value of negative equity for the 90-days-delinquency risk, the coefficients of the third and the fourth spline of original loan-to-value, the coefficients of the first and the fourth spline of log loan size for prepayment risk, the coefficients of the loan age for prepayment, the coefficient of the second spline of the credit score, and the coefficients of the first and the second spline of the debt-to-

Table 48. Results comparison for five MSAs

Explanatory Variable	Miami	Tampa	Detroit	Las Vegas
Prepayment penalty	The coefficient is within the confidence interval of Phoenix	The coefficient is within the confidence interval of Phoenix	The coefficient is outside the confidence interval of Phoenix	The coefficient is within the confidence interval of Phoenix
The value of call option	The coefficients of the second and the fourth spline are outside the confidence interval of Phoenix, and the coefficients of the first and the third spline are within the interval	The coefficient of the first spline is outside the confidence interval of Phoenix, the coefficients of other splines are within the interval	The coefficients of the second and the fourth spline are outside the confidence interval of Phoenix, and the coefficients of the first and the third spline are within the interval	The coefficients of all splines are outside the confidence interval of Phoenix
The unemployment rate	For default, the coefficients of all splines are within the confidence interval of Phoenix	For default, the coefficient of the first spline is outside the confidence interval of Phoenix, and the coefficient of the second spline is within the interval	For default, the coefficient of the second spline is outside the confidence interval of Phoenix, and the coefficient of the first spline is within the interval	For default, the coefficients of all splines are outside the confidence interval of Phoenix
	For 90-days-delinquency, the coefficient of the first spline is outside the confidence interval and the coefficient of the second spline is within the interval	For 90-days-delinquency, the coefficient of the second spline is outside the confidence interval and the coefficient of the first spline is within the interval	For 90-days-delinquency, the coefficients of all splines are outside the confidence interval	For 90-days-delinquency, the coefficients of all splines are outside the confidence interval
The value of negative equity	For default, the coefficients of all splines are within the confidence interval of Phoenix	For default, the coefficients of all splines are within the confidence interval of Phoenix	For default, the coefficients of all splines are outside the confidence interval of Phoenix	For default, the coefficients of all splines are within the confidence interval of Phoenix
	For 90-days-delinquency, the coefficient of the second spline is outside the confidence interval	For 90-days-delinquency, the coefficient of the third spline is outside the confidence interval	For 90-days-delinquency, the coefficients of the second, third and fourth spline are outside the confidence interval	For 90-days-delinquency, the coefficients of all splines are within the confidence interval
The negative equity dummy	For default, the coefficient is within the confidence interval of Phoenix	For default, the coefficient is within the confidence interval of Phoenix	For default, the coefficient is outside the confidence interval of Phoenix	For default, the coefficient is outside the confidence interval of Phoenix
	For 90-days-delinquency, the coefficient is within the confidence interval	For 90-days-delinquency, the coefficient is within the confidence interval	For 90-days-delinquency, the coefficient is outside the confidence interval	For 90-days-delinquency, the coefficient is within the confidence interval
Original LTV	For default, the coefficients of the third and the fourth spline are outside the confidence interval of Phoenix, the coefficients of other splines are within the interval	For default, the coefficient of the fourth spline is outside the confidence interval of Phoenix, the coefficients of other splines are within the interval	For default, the coefficient of the fourth spline is outside the confidence interval of Phoenix, the coefficients of other splines are within the interval	For default, the coefficients of the second, third and fourth spline are outside the confidence interval of Phoenix, the coefficients of other splines are within the interval
	For 90-days-delinquency, the coefficient of the third spline is outside the confidence interval and the coefficients of other splines are within the interval	For 90-days-delinquency, the coefficients of the third and the fourth spline are outside the confidence interval and the coefficients of other splines are within the interval	For 90-days-delinquency, the coefficients of the third and the fourth spline are outside the confidence interval and the coefficients of other splines are within the interval	For 90-days-delinquency, the coefficients of the first, third and the fourth spline are outside the confidence interval
Log loan size	For prepayment, the coefficients of the second, third and fourth spline are outside the confidence interval of Phoenix	For prepayment, the coefficients of the first and the fourth spline are outside the confidence interval of Phoenix	For prepayment, the coefficients of the third and the fourth spline are outside the confidence interval of Phoenix	For prepayment, the coefficients of the third and the fourth spline are outside the confidence interval of Phoenix
	For default, the coefficients of all splines are within the confidence interval	For default, the coefficients of all splines are within the confidence interval	For default, the coefficients of the first and the third spline are outside the confidence interval	For default, the coefficients of the second and the fourth spline are outside the confidence interval
	For 90-days-delinquency, the coefficients of all splines are within the	For 90-days-delinquency, the coefficients of all splines are within the	For 90-days-delinquency, the coefficient of the first spline is outside the	For 90-days-delinquency, the coefficients of the second and the fourth

	confidence interval	confidence interval	confidence interval	spline are outside the confidence interval
Loan age	For prepayment, the coefficients of the second and fourth spline are outside the confidence interval of Phoenix	For prepayment, the coefficients of the first, third and fourth spline are outside the confidence interval of Phoenix	For prepayment, the coefficients of all splines are outside the confidence interval of Phoenix	For prepayment, the coefficient of the first spline is outside the confidence interval of Phoenix
	For default, the coefficients of all splines are outside the confidence interval	For default, the coefficient of the second spline is outside the confidence interval	For default, the coefficients of all splines are outside the confidence interval	For default, the coefficients of the third and fourth spline are outside the confidence interval
	For 90-days-delinquency, the coefficients of all splines are outside the confidence interval	For 90-days-delinquency, the coefficient of the first spline is outside the confidence interval	For 90-days-delinquency, the coefficients of all splines are outside the confidence interval	For 90-days-delinquency, the coefficients of the first, third and fourth spline are outside the confidence interval
The credit score	For prepayment, the coefficients of the first and third spline are outside the confidence interval of Phoenix	For prepayment, the coefficient of the second spline is outside the confidence interval of Phoenix	For prepayment, the coefficients of all splines are outside the confidence interval of Phoenix	For prepayment, the coefficients of the second and fourth spline are outside the confidence interval of Phoenix
	For default, the coefficients of all splines are within the confidence interval	For default, the coefficients of all splines are within the confidence interval	For default, the coefficients of all splines are within the confidence interval	For default, the coefficients of the second and third are outside the confidence interval
	For 90-days-delinquency, coefficients of all splines are within the confidence interval	For 90-days-delinquency, coefficients of all splines are within the confidence interval	For 90-days-delinquency, coefficients of all splines are within the confidence interval	For 90-days-delinquency, the coefficient of the second spline is outside the confidence interval
Debt-to-income ratio	For prepayment, the coefficient of the first spline is outside the confidence interval of Phoenix	For prepayment, the coefficient of the fourth spline is outside the confidence interval of Phoenix	For prepayment, the coefficients of the second and fourth spline are outside the confidence of Phoenix	For prepayment, the coefficients of the first and second spline are outside the confidence of Phoenix
	For default, the coefficients of the first, second and fourth spline are outside the confidence interval	For default, the coefficients of the first and second spline are outside the confidence interval	For default, the coefficients of the third and fourth spline are outside the confidence interval	For default, the coefficients of all splines are within the confidence interval
	For 90-days-delinquency, the coefficients of the second and fourth spline are outside the confidence interval	For 90-days-delinquency, the coefficient of the first spline is outside the confidence interval	For 90-days-delinquency, the coefficient of the third spline is outside the confidence interval	For 90-days-delinquency, the coefficients of all splines are within the confidence interval
The dummy for the number of units	For prepayment, the coefficient is within the confidence interval of Phoenix	For prepayment, the coefficient is within the confidence interval of Phoenix	For prepayment, the coefficient is within the confidence interval of Phoenix	For prepayment, the coefficient is within the confidence interval of Phoenix
	For default, the coefficient is within the confidence interval	For default, the coefficient is within the confidence interval	For default, the coefficient is outside the confidence interval	For default, the coefficient is within the confidence interval
	For 90-days-delinquency, the coefficient is within the confidence interval	For 90-days-delinquency, the coefficient is within the confidence interval	For 90-days-delinquency, the coefficient is outside the confidence interval	For 90-days-delinquency, the coefficient is within the confidence interval
The confidence interval indicates the 95% confidence interval of the coefficients for Phoenix.				

income ratio for default and 90-days-delinquency are outside the 95% confidence interval of the results in Phoenix.

For Detroit, the coefficient of the indicator for prepayment penalty, the coefficient of the second and the fourth spline of the value of call option, the coefficients of the unemployment rate, the coefficients of the value of negative equity, the coefficient of the negative equity dummy, the coefficients of the third and the fourth spline of original loan-to-value, the coefficients of the first, the third and the fourth spline of log loan size, the coefficients of the loan age, the coefficients of the credit score for prepayment, and the coefficients of the second, the third and the fourth spline of the debt-to-income ratio, and the coefficient of The dummy for the number of units for default and 90-days-delinquency are outside the 95% confidence interval of the results in Phoenix. Overall, almost all of the coefficients in Detroit are outside the confident interval.

For Las Vegas, the coefficients of the value of call option, the coefficients of the unemployment rate, the coefficient of the negative equity dummy for default, the coefficients of the first, the third and the fourth spline of original loan-to-value, the coefficients of the second, the third and the fourth spline of log loan size, the coefficients of the first, the third and the fourth spline of loan age, the coefficients of the second, the third and fourth spline of the credit score, and the coefficients of the first and the second spline of the debt-to-income ratio for prepayment are outside the 95% confidence interval of the results in Phoenix.

To sum up, most of the results in Miami, Tampa, and Las Vegas support Hypotheses H1 to H8, and some of the results in these MSAs are significantly different from the results in Phoenix. In Detroit, the results partially support Hypothesis and almost all of the results are significantly different from the results in Phoenix. Therefore, in the following section, this paper presents a Wald test to show the equality of results across MSAs.

The Wald Test for the Equality of the Coefficients among Five MSAs

Follain and Giertz (2016) find evidence against the assumption of a national housing market. A related question is whether financial market for housing differs across MSAs in United State. In this section, a formal statistical test for whether people with similar characteristics across MSAs face similar outcome risks is presented. A Wald test is constructed to test the equality of the coefficients among five MSAs.

Assume the null hypothesis is:

$$H_0: \beta_{p1} = \beta_{p2} = \beta_{p3} = \beta_{p4} = \beta_{p5}$$

$$\beta_{d1} = \beta_{d2} = \beta_{d3} = \beta_{d4} = \beta_{d5}$$

In which the $\beta_{p1}$and β_{d1} are the vectors of coefficients for the prepayment and default/90-days-delinquency, respectively, for Phoenix; β_{p2} and β_{d2} are the vectors of coefficients for the prepayment and default/90-days-delinquency, respectively, for Miami; β_{p3} and β_{d3} are the vectors of coefficients for the prepayment and default/90-days-delinquency, respectively, for Tampa; $\beta_{p4}$and β_{d4} are the vectors of coefficients for the prepayment and default/90-days-delinquency, respectively, for Detroit; β_{p5} and β_{d5} are the vectors of coefficients for the prepayment and default/90-days-delinquency, respectively, for Las Vegas. The null hypothesis can also be presented as

$$H_0: \beta_{p1} - \beta_{p2} = \beta_{p2} - \beta_{p3} = \beta_{p3} - \beta_{p4} = \beta_{p4} - \beta_{p5} = 0$$

$$\beta_{d1} - \beta_{d2} = \beta_{d2} - \beta_{d3} = \beta_{d3} - \beta_{d4} = \beta_{d4} - \beta_{d5} = 0$$

A general form of the null hypothesis will be H_0: $R\beta = 0$. Assume the vector of coefficients β is constructed as $\beta = [\beta_{p1}\ \beta_{d1}\ \beta_{p2}\ \beta_{d2}\ \beta_{p3}\ \beta_{d3}\ \beta_{p4}\ \beta_{d4}\ \beta_{p5}\ \beta_{d5}]'$, to find the general form of the H_0, the R matrix is assumed to be

$$R = \begin{bmatrix} 1 & \cdots & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & -1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & -1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & -1 \end{bmatrix}$$

Then

$$R\beta = \begin{bmatrix} 1 & \cdots & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & -1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & -1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & -1 \end{bmatrix} \begin{bmatrix} \beta_{p1} \\ \beta_{d1} \\ \beta_{p2} \\ \beta_{d2} \\ \beta_{p3} \\ \beta_{d3} \\ \beta_{p4} \\ \beta_{d4} \\ \beta_{p5} \\ \beta_{d5} \end{bmatrix} = \begin{bmatrix} \beta_{p1} - \beta_{p2} \\ \beta_{d1} - \beta_{d2} \\ \beta_{p2} - \beta_{p3} \\ \beta_{d2} - \beta_{d3} \\ \beta_{p3} - \beta_{p4} \\ \beta_{d3} - \beta_{d4} \\ \beta_{p4} - \beta_{p5} \\ \beta_{d4} - \beta_{d5} \end{bmatrix}$$

Therefore, the Wald test will be

$$W = \left(R\hat{\beta} - 0\right)' \left[R\ Est\ Asy\ Var\left(\hat{\beta}\right) R'\right]^{-1} \left(R\hat{\beta} - 0\right)$$

$$= \left(R\hat{\beta}\right)' \left[R\ Est\ Asy\ Var\left(\hat{\beta}\right) R'\right]^{-1} \left(R\hat{\beta}\right)$$

The W belongs to a Chi-square distribution with the degree of freedom equal to the number of restrictions.

This paper only tests the equality of the coefficients for the important variables to determine the prepayment and default (90-days-deliqnuency) risk. The coefficients of the indicator of the prepayment penalty, the first and the last spline of the value of the call option, the first and the last spline of the value of negative equity and the negative equity dummy are tested. The results show that, $\mathrm{W} = 6065.853$ for the model using the prepayment and default as the dependent variables, and $\mathrm{W} = 3837.716$ for the model using the prepayment and 90-days-delinquency as the dependent variables. Both values are larger than ${\chi^2}_{0.05} = 36.415$ at the degree of freedom equal to 24. Therefore, the null hypothesis is rejected and the coefficients estimated for each MSA are not equal. The inequality of the coefficients among five MSAs indicates that there is not a national financial market for housing and that the financial markets for housing are heavily influenced by local characteristics. Bradley, Cutts and Liu (2015) and Lam, Dunsky, and Kelly (working paper) fail to consider this in their studies and they pool observations across locations to analyze the termination risks of single-family mortgages.

Summary

After analyzing mortgage termination behaviors in five MSAs using different dependent variables and explanatory variables, results indicate the signs of important explanatory variables are consistent with the proposed hypotheses, with some exceptions. Moreover, the results of the Wald test also support the argument that there is not a national financial market for housing and that the financial markets for housing are heavily influenced by local characteristics. The important results are summarized below.

Hypothesis 1 posits a negative relationship between the prepayment penalty indicator and the prepayment risk and is supported by the results of Phoenix, Miami, Tampa and Las Vegas. However, the effect of the indicator on the prepayment risk is insignificant in Detroit.

Hypothesis 2 states a positive relationship exists between the value of the call option and the prepayment risk and is supported in all MSAs included in this paper. The significantly positive effect of the value of the call option supports the argument that when the value of the call option is "in the money," households have more incentive to prepay their mortgages. However, the positive effect may not be significant in a few ranges of the value.

Hypothesis 3 states the monthly unemployment rate has a positive effect on the default risk and is partially supported by the results discussed in this paper. The positive effect of the unemployment rate exists in Phoenix, Tampa, Detroit and Las Vegas. When the unemployment rate is below the second quartile, the positive effect is significant in Phoenix, Miami and Las Vegas and it is significant in Tampa when the rate is above the second quartile. The effect of the unemployment rate on the 90-days-delinquency is very similar to its effect on the default.

Hypothesis 4 posits a positive relationship between negative equity and the default risk and is supported by the results of Phoenix, Miami, Tampa and Las Vegas. In Phoenix and Las Vegas, when negative equity ranges between the first quartile and the third quartile, the negative effect is significant. In Miami, when negative equity either falls below the first quartile or rises above the second quartile, the impact is significant, and in Tampa, the effect is only significant when negative equity falls below the first quartile. On the other hand, in Detroit, negative equity has a significantly positive effect on the default decision when it falls below the first quartile, and the effect becomes insignificantly negative when it rises above the first quartile. The effect of negative equity on the 90-days-delinquency is very similar to its effect on the default.

Hypothesis 5 suggests a positive coefficient for the negative equity dummy and is supported by all MSAs discussed in this paper. In Phoenix, Miami, Tampa and Las Vegas, the negative equity dummy has a significantly positive relationship with the default/90-days-delinquency risk. However, only in Detroit is the effect insignificant.

Hypothesis 6 posits that the original loan-to-value positively affects the default risk and is supported by the results. In Phoenix, Miami, Tampa and Detroit, the original loan-to-value ratio is significantly and positively correlated to the default/90-days-delinquency risk in most ranges. However, when the ratio ranges between the 75th and 95th percentile for Phoenix; between the first and the third quartile for Miami; above the third quartile for Tampa; and between the first and the third quartile for Detroit, the effect becomes insignificant. On the other hand, in Las Vegas, the effect of original loan-to-value on the default/90-days-delinquency risk is insignificant.

Hypothesis 7 states that the effect of the credit score on the default risk is negative and is supported by results in all MSAs. This paper discusses the effect of the credit score on both prepayment and default/90-days-delinquency risks, as well as the effect of the credit score only on the default/90-days-delinquency risk.

Results for Phoenix are summarized below. Generally, the credit score has a strongly positive effect on the prepayment risk (except when the credit score ranges between the second and the third quartile when the effect becomes insignificant). Comparatively, it has a strong negative impact on the default risk (except when the credit score ranges between the second and the third quartile and the effect becomes insignificant). When using the credit score only in the default risk and not in the prepayment risk, the effect of the credit score is still significantly negative, but the coefficients decrease slightly.

For Miami and Las Vegas, the models with the credit score only in the default/90-days-delinquency risk are showed in the paper because these models have relatively larger BICs. The effect of the credit score on the default/90-days-delinquency risk is negative, and the effect is significant when the credit score falls below the first quartile or rises above the second quartile.

For Tampa and Detroit, the models with the credit score in both prepayment and default risks are showed also because of the larger BIC. The effect of the credit score on the prepayment risk is positive. However, the effect is significant when the credit score rises above the third quartile for Tampa and falls below the first quartile or ranges between the second and the third quartile for Detroit. Comparatively, the effect of the credit score on the default/90-days-delinquency risk is significantly negative, and the effect is insignificant for the credit score ranges between the second and the third quartile.

Hypothesis 8 posits that the relationship between the debt-to-income ratio and the default risk is positive and is partially supported by the results. For Phoenix, Miami, Tampa and Las Vegas the effect of the debt-to-income ratio is not consistent for the prepayment risk, and it has a insignificantly positive relationship with the default/90-days-delinquency risk. For Detroit, the effect of the debt-to-income ratio is inconsistent for both the prepayment and default/90-days-delinquency risks.

In Phoenix, Miami, Tampa and Las Vegas, the effect of The dummy for the number of units is negative on both the prepayment and default/90-days-delinquency risks. This indicates mortgages covering more than one house unit are less likely to prepay or default. However, in Detroit, The dummy for the number of units has a significantly negative impact on the prepayment risk and a significantly positive impact on the default/90-days-delinquency risk.

There, mortgages covering more than one house unit are more likely to default/delinquent than to prepay.

Generally, the younger mortgages, which stay on the market less than one or two years, are more likely to be either prepaid or defaulted/delinquent. But the older mortgages, which survived in the market for a long time, are more likely to be prepaid than defaulted or delinquent. However, the results of this study indicate the effect of the log loan size is highly sensitive to which dependent variables and explanatory variables are being used.

This study also constructs a Wald test to test the equality of the coefficients among five MSAs and the result shows that the coefficients among MSAs are significantly different. This result indicates that there is not a national financial market for housing and that the financial markets for housing are heavily influenced by local characteristics. Bradley, Cutts and Liu (2015) and Lam, Dunsky, and Kelly (working paper) do not consider this in their studies and they pool observations across locations to analyze the termination risks of single-family mortgages.

Chapter 2

Introduction

In earlier studies, the proportional hazards model has been widely used to analyze the prepayment or default risks of mortgages. Green and Shoven (1986), Schwart and Torous (1989), Quigley and Van Order (1990), Follain, Ondrich and Sinha (1995) use this model to analyze the prepayment risk; Quigley and Van Order (1991, 1995) wrote a series of papers on implementing the method to estimate the default risk. The basic likelihood function based on the proportional hazards model is showed in the following:

$$l(\beta) = \prod_{i=1}^{N}\{[1-\exp(-\exp(z_i(t)\beta))]^{\delta_i} \times \prod_{s=1}^{t-1}\exp(-\exp(z_i(s)\beta))\} \qquad 2.1$$

where δ_i equal 1 when individual i terminates the status, $z_i(t)$ is the individual specific characteristics, and β is a vector of coefficients to be estimated.

In more recent studies, the competing risks model has become a popular methodology to analyze the prepayment and default risks simultaneously in the mortgage market. Among different frameworks used to construct the competing risks model, the proportional hazards model is widely used. One way to directly construct the competing risks model using the 2.1 framework is

$$l(\beta) = \prod_{i=1}^{N}\prod_{s=1}^{t}\left[(1-\exp(-\exp\left(z_p(s)\beta_p\right)))\exp(-\exp(z_d(s)\beta_d))\right]^{\delta_p}\left[(1 - \exp(-\exp(z_d(s)\beta_d)))\exp(-\exp\left(z_p(s)\beta_p\right))\right]^{\delta_p}\exp\left(-\left(\exp(z_d(s)\beta_d) + \exp\left(z_p(s)\beta_p\right)\right)\right)^{1-\delta_p-\delta_d} \qquad 2.2$$

An important assumption for model 2.2 is that the time in the analysis is a discrete time. When a default/prepayment occurs in the time period $t-1$ to t, regardless of whether the mortgage is continued or prepaid/defaulted after time t, the mortgage is treated as default/prepayment and is removed from the sample. One significant weakness of this assumption is that it does not consider the time period between the default/prepayment point and the time point t. Therefore, model 2.2 could lead to estimation bias. To include the time period between the default/prepayment point and the time point t in the calculation, Sueyoshi (1992) introduces a methodology based on double integrals to construct a more accurate competing risks model. The concept of the methodology is widely used in later studies. Deng, Quigley and Van

Order (1996; 2000) and Deng (1997) are examples of implementing the method to analyze the termination risks of the thirty-year, fixed-rate, single-family mortgages, and Wenyi Huang and Jan Ondrich (2002) is an example of analyzing the termination risks of the multifamily mortgages.

Although the model based on Sueyoshi's methodology is widely used in the literature, the process of constructing the competing risks model is ambiguous. This paper clearly presents the process of calculating this model based on the proportional hazards model using Sueyoshi's method and then implements the model to analyze the termination risks of single-family mortgages in Phoenix.

Ding, Tian, Yu and Guo (2012) construct a new model based on a class of transformation survival models to analyze the risk of bankruptcy. The model is controlled by a transformation parameter c. When $c = 0$, the model is the proportional hazards model and when $c > 0$, the model transfers to different forms. An important argument made by the authors is that the proportional hazards model is not the best model to analyze the risk of bankruptcy. Therefore, a question is raised by this argument: whether the proportional hazards model is the best model to analyze the default/prepayment risk of single-family mortgages, and if it is not which model will be the best?

To answer these questions, this paper constructs a new competing risks model based on a class of transformation survival models and Sueyoshi's method and uses it to analyze the termination risks of the single-family mortgages in Phoenix. The model is controlled by the transformation parameters c_p (for the prepayment risk) and c_d (for the default risk). When $c_p = 0$ and $c_d = 0$, it is the competing risks model based on the proportional hazards model, and when

$c_p > 0$and $c_d > 0$, its framework is changed according to the value of c_p and c_d. The results show that the proportional hazards framework is the best model to estimate the prepayment risk, but it is not the best model to estimate the default/90-days-delinquency risk.

The results of both models support the important arguments made in Chapter 1. The prepayment penalty has a significantly negative effect on the prepayment decision. The effect of the value of the call option on the prepayment risk is significantly positive. In addition, the unemployment rate has a positive relationship with the default/90-days-delinquency. Moreover, the value of negative equity, the negative equity dummy and the original loan-to-value generally are positively related to the default/90-days-delinquency. As expected, the credit score has a strong positive effect on the prepayment; comparatively, it has a strong negative impact on the default. The debt-to-income ratio does not appear to affect prepayment, but it has a positive effect on the default/90-days-delinquency. Comparing the coefficients estimated by three competing risks models, the coefficients estimated by the model based on the Sueyoshi's proportional hazards are insignificantly distinguishable in sign from those estimated by the model based on multinomial logit. Moreover, the coefficients estimated by the model based on a class of transformation survival models are significantly different from those estimated by the other two models.

The remaining information in this chapter is separated into three sections. In section II, literature reviews and the detail process of constructing the competing risks model is introduced. In section III, the dataset and results are clearly discussed. The summary is offered in section IV.

Literature Review

As noted above, the proportional hazards model is the most popular method used to analyze the prepayment or default risks of mortgages in previous studies.

As the first researcher who introduced the proportional hazards model into the discrete time survival analysis, Cox (1972, 1975) discusses a method to estimate the coefficients in the proportional hazards model by a partial likelihood function. However, the method is not computationally feasible with large data sets that contain many failure times or with grouped survival data.

Prentice and Gloeckler (1978)[11] develop a maximum likelihood estimation to estimate the coefficients in the proportional hazards model and implement it to analyze the survival rate of 11,442 breast cancer patients. The detail process of constructing the model follows.

Let T_i be a random variable representing the failure time of individual i. Let $Z_i(t)$ be a vector of time-dependent characteristics for individual i. Finally, let δ_i equal 1 when individual i terminate the status. Then the hazard for individual i at time t, $\lambda_i(t)$, is defined by the equation

$$\lim_{h \to 0^+} \frac{prob[t + h > T_i \geq t | T_i \geq t]}{h} = \lambda_i(t)$$

The hazard is parameterized using the Cox proportional hazards form

$$\lambda_i(t) = \lambda_0(t) \exp(Z_i(t)\beta) \tag{2.3}$$

where $\lambda_0(t)$ is the baseline hazard at time t.

The probability of an individual surviving until time t+1 conditional on that it has survived until t is given by

[11] The model introduced by Prentice and Gloeckler (1978) is a general model, which can be implemented in any survival analysis. However, this paper mainly focuses on using it to analyze the termination risk of single-family mortgages.

$$P[T_i \geq t+1 | T_i \geq t] = \exp\left(-\int_t^{t+1} \lambda_i(u)du\right)$$

$$= \exp\left(-\exp(Z_i(t)\beta)\int_t^{t+1} \lambda_0(u)du\right)$$

$$= \exp(-\exp(Z_i(t)\beta + \gamma(t)))$$

$$= \exp(-\exp(z_i(t)\beta)) \qquad 2.4$$

where $\gamma(t) = \ln(\int_t^{t+1} \lambda_0(u)du)$, $z_i(t)$ contains $Z_i(t)$ and the baseline, and the β in the last line of 2.2 contains $\gamma(t)$.

The likelihood function for a sample of N individuals will be

$$l(\beta) = \prod_{i=1}^{N} \{[1 - \exp(-\exp(z_i(t)\beta))]^{\delta_i} \times \prod_{s=1}^{t-1} \exp(-\exp(z_i(s)\beta))\} \qquad 2.5$$

and the log-likelihood function is

$$L(\beta) = \sum_{i-1}^{N} \{\delta_i log[1 - \exp(-\exp(z_i(t)\beta))] - \sum_{s=1}^{t-1} \exp(z_i(s)\beta)\} \qquad 2.6$$

where $\mathrm{t} = \min(\mathrm{int}(\mathrm{T_i}), \mathrm{C_i})$ and $\mathrm{C_i}$ is the censoring time.

Meyer (1990) points out that the Prentice and Gloeckler approach makes no assumptions about the baseline hazard. By reconstructing the function 2.3, Meyer includes unobserved heterogeneity in the following the model:

$$\lambda_i(t) = \theta_i \lambda_0(t) \exp(Z_i(t)\beta) \qquad 2.7$$

where θ_i is a random variable that is assumed to be independent of $Z_i(t)$. Then the log-likelihood function 2.6 becomes

$$L(\beta,\gamma,\theta)=\sum_{i-1}^{N}\{log[\int \exp\left[-\theta\sum_{s=0}^{t-1}\exp\left(Z_i(s)\beta+\gamma(s)\right)\right]du(\theta)-\delta_i\int \exp\left[-\theta\sum_{s=0}^{t}\exp\left(Z_i(s)\beta+\gamma(s)\right)\right]du(\theta)]\} \quad 2.8$$

$u(\theta)$ is the distribution of θ and the commonly used distribution is the gamma distribution with mean one and variance σ^2[12]. Function 2.6 is the fundamental approach used in the previous studies to analyze either the prepayment or default risks in the mortgages market. Studies by Green and Shoven (1986), Schwart and Torous (1989) and Quigley and Van Order (1990) implement the method to analyze the prepayment risk, and studies by Quigley and Van Order (1991, 1995) use the method to analyze the default risk. When the unobserved heterogeneity is considered in the analysis, function 2.8 is applied. The paper by Follain, Ondrich and Sinha (1995) is one of the examples that uses the model to analyze the prepayment risk of the multifamily mortgages.

The methodology by Prentice and Gloeckler (1978) and Meyer (1990) is only useful for analyzing the event with a single failure type. When there are two or more failure types involved into the analysis, which means there are at least three choices (continue the event, terminate the event with prepayment, and terminate the event with default) to be chosen in each time period, the competing risks model should be applied.

Sueyoshi (1992) develops the proportional hazards model into a competing risks model by writing the likelihood of an event that has not terminated at the end of the sample period t as:

$$L_t^C(\beta)=\int_{\alpha_t^1}^{\infty}\int_{\alpha_t^2}^{\infty}f(u_1,u_2)du_1du_2 \quad 2.9$$

---

[12] If the distribution for θ is gamma with mean one and variance σ^2, the model 2.8 becomes:

$$L(\beta,\gamma,\sigma^2)=\sum_{i-1}^{N}\log\{\left[1+\sigma^2\sum_{s=0}^{t-1}\exp\left(Z_i(s)\beta+\gamma(s)\right)\right]^{-\sigma^{-2}}-\delta_i\left[1+\sigma^2\sum_{s=0}^{t}\exp\left(Z_i(s)\beta+\gamma(s)\right)\right]^{-\sigma^{-2}}\} \quad 2.8$$

The likelihood of an event terminating with prepayment is

$$L_t^P(\beta) = \int_{\alpha_{t-1}^1}^{\alpha_t^1} \int_{g(u_1)}^{\infty} f(u_1, u_2) du_1 du_2 \qquad 2.10$$

in which $g(u_1) = \alpha_{t-1}^2 + (u_1 - \alpha_{t-1}^1)\frac{\alpha_t^2 - \alpha_{t-1}^2}{\alpha_t^1 - \alpha_{t-1}^1}$.

Finally, the likelihood of an event terminating with default is

$$L_t^D(\beta) = \int_{\alpha_{t-1}^2}^{\alpha_t^2} \int_{g(u_2)}^{\infty} f(u_1, u_2) du_1 du_2 \qquad 2.11$$

in which

$$g(u_2) = \alpha_{t-1}^1 + (u_2 - \alpha_{t-1}^2)\frac{\alpha_t^1 - \alpha_{t-1}^1}{\alpha_t^2 - \alpha_{t-1}^2}$$

$$f(u) = \exp(-u)$$

$$\alpha_t^1 = \sum_{s=1}^{t} \exp\left(z_{pi}(s)\beta_p\right)$$

$$\alpha_t^2 = \sum_{s=1}^{t} \exp(z_{di}(s)\beta_d)$$

$$\alpha_{t-1}^1 = \sum_{s=1}^{t-1} \exp\left(z_{pi}(s)\beta_p\right)$$

$$\alpha_{t-1}^2 = \sum_{s=1}^{t-1} \exp(z_{di}(s)\beta_d)$$

The competing risks model based on the proportional hazards model will be

$$L(\beta)=\sum_{i=1}^{N}\sum_{s=1}^{t}\left\{\delta_p\left[-\ln\left(1+\frac{\exp(z_{di}(s)\beta_d)}{\exp\left(z_{pi}(s)\beta_p\right)}\right)+\ln\left(1-\exp\left(-\left(\exp\left(z_{pi}(s)\beta_p\right)+\exp(z_{di}(s)\beta_d)\right)\right)\right)\right]\right.$$

$$+\delta_d\left[-\ln\left(1+\frac{\exp\left(z_{pi}(s)\beta_p\right)}{\exp(z_{di}(s)\beta_d)}\right)+\ln\left(1-\exp\left(-\left(\exp(z_{di}(s)\beta_d)+\exp\left(z_{pi}(s)\beta_p\right)\right)\right)\right)\right]$$

$$\left.+\delta_c\left[-\exp\left(z_{pi}(s)\beta_p\right)-\exp\left(z_{di}(s)\beta_d\right)\right]\right\} \qquad 2.12$$

in which $z_{pi}(t), z_{di}(t)$ are vectors of time-dependent explanatory variables for event i to be terminated with prepayment and default, respectively. β_p, β_d are coefficients to be estimated. Moreover, δ_p is the indicator for whether prepayment is chosen, δ_d is the indicator for whether default is chosen and δ_c is the indicator for whether the event continues without termination. The general calculation process is shown in Appendix B.

The basic structure of the above competing risks model is frequently used in more recent empirical studies to analyze the competing risks of prepayment and default. The papers by Deng, Quigley and Van Order (1996) and Deng (1997) are examples using the method to analyze thirty-year, fixed-rate, single-family mortgages issued from 1976 to 1983. As previously mentioned, Deng, Quigley and Van Order (2000), and Wenyi Huang and Jan Ondrich (2002) incorporate the unobserved heterogeneity into the competing risks model, with the former using it to analyze the single-family mortgages and the latter using it to analyze the multifamily mortgages.

As mentioned above, a study by Ding, Tian, Yu and Guo (2012) [13] introduces a new model based on a class of discrete transformation survival models. This method gives the estimation more flexibility. The detail model is listed in the following.

Let $S(t) = P(T \geq t), t = 1, \ldots t_k$ be the survival function based on the explanatory variables $Z(\ t)$. The conditional probability that individual i fails at time t given it has survived at time $t-1$ is

$$p_{it} = P(T = t | T \geq t - 1) \tag{2.13}$$

Also, let

$$G\left[-log\frac{S(t)}{S(t-1)}\right] = \exp(Z_i(t)\beta)\, G\left[-log\frac{S_0(t)}{S_0(t-1)}\right] \tag{2.14}$$

where G was a strictly increasing transformation function with $G(0) = 0$ and $G(\infty) = \infty$. $S_0(.)$is the baseline survival function when $Z(\ t) = 0$.

When transformation function G belongs to the family

$$G(x) = \begin{cases} \frac{1}{c}[\exp(cx) - 1], c > 0 \\ x, \qquad\qquad\qquad c = 0 \end{cases} \tag{2.15}$$

the class of discrete-time transformation survival models will become

$$p_{it} = \begin{cases} 1 - \dfrac{1}{[1 + cexp(\alpha + Z(t)\beta)]^{\frac{1}{c}}}, & c > 0 \\ 1 - \exp[-\exp(\alpha + Z(t)\beta)], & c = 0 \end{cases} \tag{2.16}$$

[13] The model introduced by Ding, Tian, Yu and Guo (2012) is used to analyze the firm bankruptcy; therefore, their assumption is based on whether a firm files for bankruptcy at time t. However, this paper is based on a more general assumption to make a more general model.

Note that when $c = 0$, the proposed discrete-time transformation survival model is equivalent to the traditional proportional hazards model. The likelihood function for the proposed discrete transformation survival model for a sample of N is

$$l(\beta) = \prod_{i=1}^{N} \prod_{s=1}^{t} p_{is}{}^{\delta_i} (1 - p_{is})^{(1-\delta_i)} \qquad 2.17$$

in which $\delta_i = 1$ when individual i fails at time t.

In the paper by Ding *et al.* (2012), the model is used to analyze the risk of firm bankruptcy based on a dataset that contains all publicly traded companies in the United States between 1981 and 2006. The estimated coefficients change according to the different values of the *c* being controlled. Moreover, the results show that when *c* is around 10, the model reaches the maximum log likelihood. This suggests that the traditional proportional hazards model is not the best framework to estimate the risk of the bankruptcy.

When two different termination types (such as prepayment and default) are involved in the survival analysis, a competing risks model is developed based on Ding *et al.* (2012) and Sueyoshi (1992). The log-likelihood function for a sample of N is

$$\begin{aligned} L(\beta) = \sum_{i=1}^{N} \sum_{s=1}^{t} \Bigg\{ & \delta_p \Bigg[-\ln\left(1 + \frac{G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}\right) \\ & + \ln\Big(1 - \exp\Big(-\big(G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)] + G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\big)\Big)\Big)\Bigg] \\ & + \delta_d \Bigg[-\ln\left(1 + \frac{G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}{G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]}\right) \\ & + \ln\Big(1 - \exp\Big(-\big(G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)] + G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\big)\Big)\Big)\Bigg] \\ & + \delta_c\big[- G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)] - G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\big] \Bigg\} \end{aligned} \qquad 2.18$$

The transformation function G is given by:

$$G_p(x) = \begin{cases} \frac{1}{c_p}\left[\exp\left(c_p x\right) - 1\right], c_p > 0 \\ x, \qquad\qquad\qquad\quad c_p = 0 \end{cases}$$

$$G_d(x) = \begin{cases} \frac{1}{c_d}\left[\exp(c_d x) - 1\right], c_d > 0 \\ x, \qquad\qquad\qquad\quad c_d = 0 \end{cases}$$

Hence, when $G_p(x) = x$ and $G_d(x) = x$, the log-likelihood function for the sample is the same as the function 2.12. When $G_p(x) = \frac{1}{c_p}\left[\exp\left(c_p x\right) - 1\right]$ and $G_d(x) = \frac{1}{c_d}[\exp(c_d x) - 1]$, the log-likelihood function for the sample is:

$$L(\beta) = \sum_{i=1}^{N}\sum_{s=1}^{t}\left\{\delta_p\left[-\ln\left(1 + \frac{\frac{1}{c_d}\log(1 + c_d\exp(z_{di}(s)\beta_d))}{\frac{1}{c_p}\log(1 + c_p\exp\left(z_{pi}(s)\beta_p\right))}\right)\right.\right.$$

$$\left. + \ln\left(1 - [1 + c_p\exp\left(z_{pi}(s)\beta_p\right)]^{-\frac{1}{c_p}}[1 + c_d\exp(z_{di}(s)\beta_d)]^{-\frac{1}{c_d}}\right)\right]$$

$$+ \delta_d\left[-\ln\left(1 + \frac{\frac{1}{c_p}\log(1 + c_p\exp\left(z_{pi}(s)\beta_p\right))}{\frac{1}{c_d}\log(1 + c_d\exp(z_{di}(s)\beta_d))}\right)\right.$$

$$\left. + \ln\left(1 - [1 + c_p\exp\left(z_{pi}(s)\beta_p\right)]^{-\frac{1}{c_p}}[1 + c_d\exp(z_{di}(s)\beta_d)]^{-\frac{1}{c_d}}\right)\right]$$

$$\left. + \delta_c\left[-\frac{1}{c_p}\log(1 + c_p\exp\left(z_{pi}(s)\beta_p\right)) - \frac{1}{c_d}\log(1 + c_d\exp(z_{di}(s)\beta_d))\right]\right\} \qquad 2.19$$

Moreover, when $G_p(x) = x$ and when $G_d(x) = \frac{1}{c_d}[\exp(c_d x) - 1]$, the log-likelihood function for the sample is:

$$L(\beta) = \sum_{i=1}^{N} \sum_{s=1}^{t} \Bigg\{ \delta_p \left[-\ln\left(1 + \frac{\frac{1}{c_d}\log(1 + c_d \exp(z_{di}(s)\beta_d))}{\exp\left(z_{pi}(s)\beta_p\right)} \right) \right.$$

$$\left. + \ln\left(1 - \exp(-\exp\left(z_{pi}(s)\beta_p\right) - \frac{1}{c_d}\ln(1 + c_d \exp(z_{di}(s)\beta_d)))\right) \right]$$

$$+ \delta_d \left[-\ln\left(1 + \frac{c_d \exp\left(z_{pi}(s)\beta_p\right)}{\log(1 + c_d \exp(z_{di}(s)\beta_d))} \right) \right.$$

$$\left. + \ln\left(1 - \exp(-\exp\left(z_{pi}(s)\beta_p\right) - \frac{1}{c_d}\ln(1 + c_d \exp(z_{di}(s)\beta_d)))\right) \right]$$

$$+ \delta_c \left[-\exp\left(z_{pi}(s)\beta_p\right) - \frac{1}{c_d}\log(1 + c_d \exp(z_{di}(s)\beta_d)) \right] \Bigg\} \qquad 2.20$$

The detailed calculation process is in Appendix C.

Results

Models 2.12 and 2.19 are first used to analyze the competing risks of the prepayment and default of the single-family mortgages in Phoenix. The results of model 2.19 show that when $c_p = 0.00003$ and $c_d = 268.49$, the maximum log-likelihood is reached for monthly data. The trend of the log-likelihood based on the change of c_p is shown in figures 24 and 25. It presents that with the value of c_p decreased to 0^+, the value of the log-likelihood continues to increase. This suggests the proportional hazards framework is the best option to estimate the prepayment risk. Moreover, the trend of the log-likelihood based on the change of c_d is showed in figures 26 and 27. It presents that with the value of c_d increased, the value of the log-likelihood first increases to a maximum point and then continues to decrease. This suggests that an optimal value of c_d needs to be estimated to get the maximum log likelihood. Therefore, model 2.20 is used to estimate the competing risks of prepayment and default and the results are showed in the following.

[insert Figures 24 through 27 here]

This section clearly explains the results of models 2.12 and 2.20. Then the results are compared with the 95 percent confident interval of the coefficients estimated by the competing risks model based on the multinomial logit. For each model, two sets of dependent variables are analyzed, the prepayment and the default being the first, and the prepayment and the 90-days-delinquency being the second. For both groups of dependent variables, the indicator for a prepayment penalty and the value of the call option are used to explain the prepayment risk; the unemployment rate, negative equity, a dummy for negative equity and original loan-to-value are used to explain the default/90-days-delinquency risk; the credit score, the debt-to-income ratio, log mortgage size, current loan age in months, a dummy for house unit and time period dummies are used to explain both risks. Moreover, the dataset used in this chapter is the dataset with the modified mortgages being right censored one month before the modification date.

Model 1: The competing risks model based on Sueyoshi's proportional hazards model

Table 50 shows the results of the competing risks of prepayment and default/90-days-delinquency estimated by model 2.12. The results of the model using prepayment and default as dependent variables are first explained in detail. The difference of the results of the model using prepayment and 90-days-delinquency as dependent variables is then discussed.

[insert Tables 49 through 50 here]

Figure 24. log-likelihood values for monthly prepayments based on the change of cp

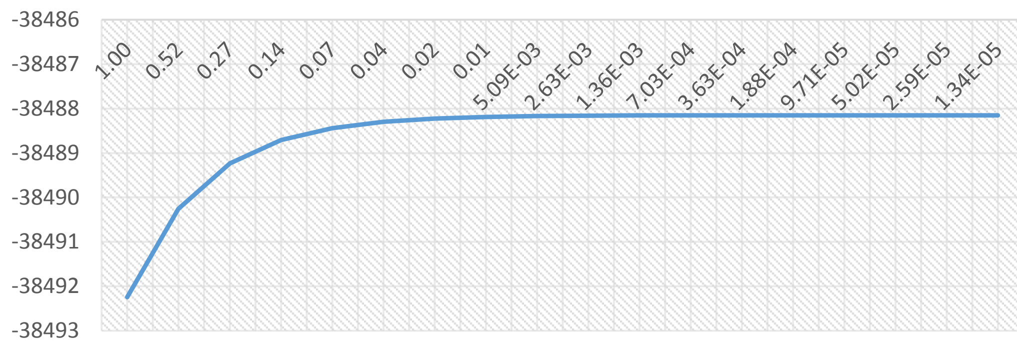


Figure25. log-likelihood values for yearly prepayments based on the change of cp

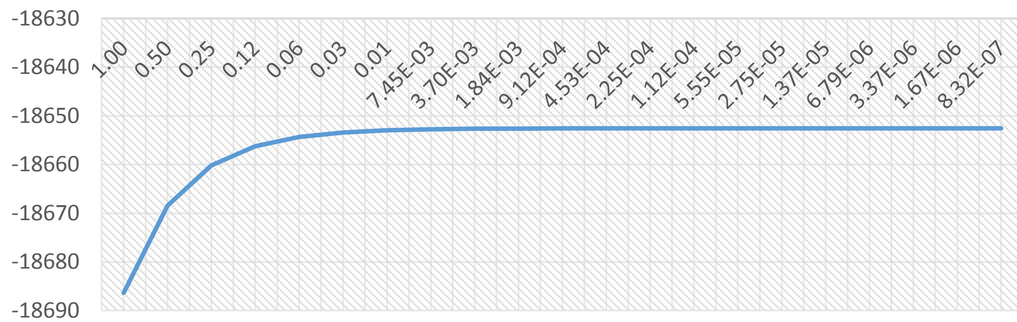


Figure 26. log-likelihood values for monthly defaults based on the change of cd

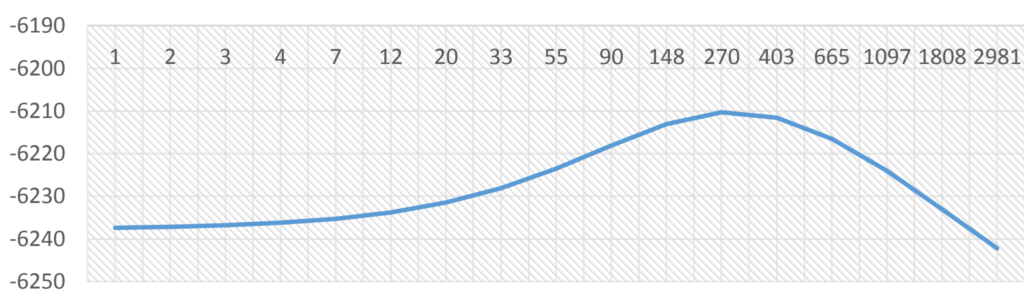


Figure 27. log-likelihood values for yearly defaults based on the change of cd

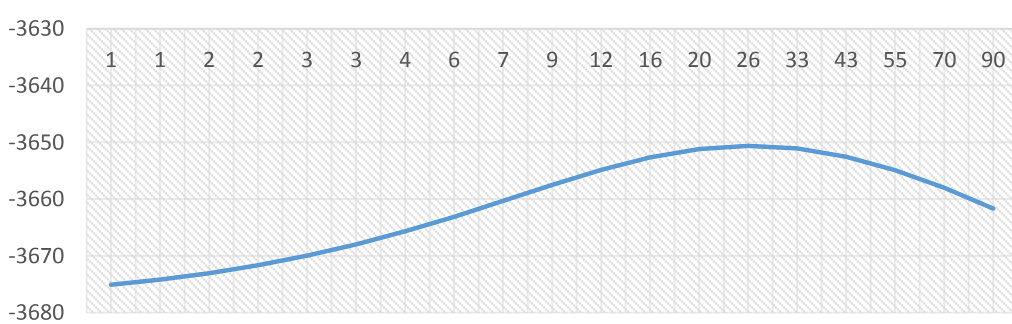


Table 49. The monthly results for Phoenix based on Meyer's proportional hazards model

	Dependent Variables			
	Prepayment (Default is right censored)	Default (Prepayment is right censored)	Prepayment (90-days-delinquency is right censored)	90-days-delinqeuncy (Prepayment is right censored)
Prepayment penalty	-0.803 (0.230)	--	-0.817 (0.231)	--
Call option variables				
Call option part 1	5.347 (0.879)	--	5.361 (0.887)	--
Call option part 2	7.403 (0.883)	--	7.270 (0.902)	--
Call option part 3	7.837 (0.663)	--	8.300 (0.671)	--
Call option part 4	1.867 (0.418)	--	2.146 (0.414)	--
The unemployment rate variables				
The unemployment rate part 1	--	0.272 (0.104)	--	0.308 (0.076)
The unemployment rate part 2	--	0.112 (0.058)	--	0.190 (0.051)
Negative equity variables				
Negative equity part 1	--	0.012 (0.014)	--	0.012 (0.014)
Negative equity part 2	--	0.018 (0.009)	--	0.019 (0.008)
Negative equity part 3	--	0.013 (0.005)	--	0.017 (0.005)
Negative equity part 4	--	-0.000 (0.003)	--	-0.000 (0.003)
The negative equity dummy	--	0.935 (0.211)	--	0.780 (0.197)
Original loan-to-value variables				
Original LTV part 1	--	3.192 (1.170)	--	2.278 (0.910)
Original LTV part 2	--	1.738 (1.222)	--	0.509 (1.098)
Original LTV part 3	--	-6.188 (5.325)	--	-7.088 (4.952)
Original LTV part 4	--	3.334 (0.749)	--	3.320 (0.694)
Original LTV part 5	--	2.371 (4.506)	--	2.423 (4.188)
Log loan size variables				
Log loan size part 1	0.746 (0.077)	0.348 (0.370)	0.778 (0.078)	0.135 (0.312)
Log loan size part 2	0.829 (0.122)	-0.453 (0.425)	0.824 (0.122)	0.012 (0.383)
Log loan size part 3	0.206 (0.130)	-0.124 (0.373)	0.218 (0.132)	-0.526 (0.345)

Log loan size part 4	0.979	-0.023	0.971	-0.049
	(0.093)	(0.267)	(0.093)	(0.244)
Loan age variables				
Loan age part 1	0.139	0.418	0.138	0.240
	(0.007)	(0.116)	(0.008)	(0.048)
Loan age part 2	0.002	0.077	0.003	0.040
	(0.003)	(0.017)	(0.004)	(0.014)
Loan age part 3	-0.002	-0.005	-0.002	-0.007
	(0.002)	(0.006)	(0.002)	(0.005)
Loan age part 4	0.006	-0.006	0.006	-0.005
	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables				
The credit score part 1	1.771	-7.898	1.335	-8.340
	(0.670)	(1.470)	(0.663)	(1.269)
The credit score part 2	3.440	-5.847	3.467	-8.840
	(0.923)	(2.471)	(0.923)	(2.287)
The credit score part 3	0.936	-2.685	0.862	-5.026
	(1.212)	(3.892)	(1.248)	(3.864)
The credit score part 4	6.864	-20.699	6.608	-21.205
	(1.637)	(6.729)	(1.643)	(6.618)
The debt-to-income ratio variables				
DTI part 1	-0.848	1.156	-0.859	2.351
	(0.369)	(1.557)	(0.369)	(1.530)
DTI part 2	0.927	2.772	1.020	3.430
	(0.495)	(1.574)	(0.495)	(1.465)
DTI part 3	-0.664	0.724	-0.679	1.097
	(0.539)	(1.440)	(0.541)	(1.316)
DTI part 4	0.326	0.645	0.354	0.543
	(0.344)	(0.756)	(0.345)	(0.693)
The dummy for the number of units	-0.541	-0.525	-0.599	-0.028
	(0.224)	(0.709)	(0.230)	(0.503)
Time period variables				
Time period part 1	0.022	-0.004	0.022	0.002
	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	-0.037	0.035	-0.036	0.001
	(0.001)	(0.010)	(0.001)	(0.009)
Time period part 3	0.017	0.005	0.016	0.008
	(0.001)	(0.008)	(0.001)	(0.008)
Prepayment dummy	-15.852	--	-15.915	--
	(0.992)		(0.991)	
Default dummy	--	-15.063	--	-9.479
		(4.506)		(3.663)
	Number of mortgages 12,734		Number of mortgages 12,698	

Table 50. Monthly results for Phoenix: the competing risks models based on Sueyoshi's proportional hazards model

	The competing risks of the prepayment and the default		The competing risks of the prepayment and the 90-days-delinqeuncy	
	Prepayment	Default	Prepayment	90-days-delinqeuncy
Prepayment penalty	-0.803 (0.230)	--	-0.809 (0.230)	--
Call option variables				
Call option part 1	5.342 (0.879)	--	5.377 (0.887)	--
Call option part 2	7.394 (0.883)	--	7.238 (0.902)	--
Call option part 3	7.817 (0.663)	--	8.283 (0.671)	--
Call option part 4	1.838 (0.418)	--	2.123 (0.414)	--
The unemployment rate variables				
The unemployment rate part 1	--	0.269 (0.104)	--	0.306 (0.076)
The unemployment rate part 2	--	0.112 (0.058)	--	0.189 (0.051)
Negative equity variables				
Negative equity part 1	--	0.012 (0.014)	--	0.012 (0.014)
Negative equity part 2	--	0.018 (0.009)	--	0.020 (0.008)
Negative equity part 3	--	0.013 (0.005)	--	0.017 (0.005)
Negative equity part 4	--	-0.000 (0.003)	--	-0.000 (0.003)
The negative equity dummy	--	0.934 (0.211)	--	0.780 (0.197)
Original loan-to-value variables				
Original LTV part 1	--	3.226 (1.171)	--	2.285 (0.909)
Original LTV part 2	--	1.725 (1.222)	--	0.495 (1.098)
Original LTV part 3	--	-6.195 (5.324)	--	-7.021 (4.954)
Original LTV part 4	--	3.326 (0.749)	--	3.313 (0.694)
Original LTV part 5	--	2.375 (4.508)	--	2.419 (4.188)
Log loan size variables				
Log loan size part 1	0.745 (0.078)	0.340 (0.370)	0.775 (0.078)	0.137 (0.312)
Log loan size part 2	0.828 (0.122)	-0.460 (0.425)	0.825 (0.122)	-0.015 (0.382)
Log loan size part 3	0.205 (0.130)	-0.127 (0.373)	0.215 (0.132)	-0.520 (0.345)

Log loan size part 4	0.979	-0.033	0.970	-0.062
	(0.093)	(0.267)	(0.093)	(0.244)
Loan age variables				
Loan age part 1	0.139	0.417	0.138	0.239
	(0.007)	(0.116)	(0.008)	(0.048)
Loan age part 2	0.002	0.077	0.003	0.039
	(0.003)	(0.017)	(0.004)	(0.014)
Loan age part 3	-0.002	-0.006	-0.002	-0.007
	(0.002)	(0.006)	(0.002)	(0.005)
Loan age part 4	0.006	-0.006	0.006	-0.005
	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables				
The credit score part 1	1.784	-7.901	1.353	-8.334
	(0.670)	(1.469)	(0.663)	(1.269)
The credit score part 2	3.453	-5.916	3.449	-8.913
	(0.923)	(2.471)	(0.923)	(2.288)
The credit score part 3	0.945	-2.736	0.942	-4.977
	(1.212)	(3.894)	(1.248)	(3.864)
The credit score part 4	6.870	-20.730	6.470	-21.226
	(1.637)	(6.733)	(1.644)	(6.615)
The debt-to-income ratio variables				
DTI part 1	-0.846	1.170	-0.862	2.407
	(0.369)	(1.557)	(0.369)	(1.531)
DTI part 2	0.925	2.750	1.012	3.331
	(0.495)	(1.574)	(0.496)	(1.465)
DTI part 3	-0.664	0.747	-0.684	1.191
	(0.539)	(1.440)	(0.541)	(1.317)
DTI part 4	0.325	0.633	0.358	0.523
	(0.344)	(0.756)	(0.345)	(0.693)
The dummy for the number of units	-0.539	-0.522	-0.602	-0.031
	(0.224)	(0.708)	(0.230)	(0.504)
Time period variables				
Time period part 1	0.022	-0.004	0.022	0.002
	(0.002)	(0.003)	(0.002)	(0.003)
Time period part 2	-0.037	0.035	-0.036	0.001
	(0.001)	(0.010)	(0.001)	(0.009)
Time period part 3	0.017	0.004	0.016	0.007
	(0.001)	(0.008)	(0.001)	(0.008)
Prepayment dummy	-15.854	--	-15.887	--
	(0.992)		(0.991)	
Default dummy	--	-14.984	--	-9.527
		(4.505)		(3.665)
	Number of mortgages 12,734		Number of mortgages 12,698	

The effect of the indicator for a prepayment penalty is negative. The odds for mortgages with a prepayment penalty are only 0.448 times as high as the same odds for mortgages without a prepayment penalty.

The effect of the value of the call option is always positive and the significantly positive effect supports the argument that when the value of the call option is "in the money," households have more incentive to prepay their mortgages. Because the range for any spline segment is less than one, it is meaningless to discuss a unit increase in the call option. Therefore, the odds ratios for the value of the call option are not discussed. Instead, this paper directly looks at the coefficients. This holds true throughout the paper whenever any explanatory variable has a spline segment less than one. When the value of the call option is below -0.025, the coefficient is 5.342 (the standard error is 0.879). When the value of the call option ranges between -0.025 and 0.037, the coefficient increases to 7.394 (the standard error is 0.883). When the value of the call option ranges between 0.037 and 0.102, the coefficient again increases to 7.817 (the standard error is 0.663). However, when the value of the call option is above 0.102, the coefficient drops to 1.838 (the standard error is 0.418).

The positive effect of the monthly unemployment rate indicates that the default risk is higher in the months with the higher unemployment rate. When the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.309 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity increase by 1.119 times with a 1 percent increase in the unemployment rate.

The results of negative equity strongly support the argument that when the put option is "in the money," households have more incentive to default their mortgages. When negative equity is below 18.366, the odds of default relative to continuity increase by 1.012 times with a 1 unit increase in negative equity. When negative equity ranges between 18.366 and 37.804, the odds of default relative to continuity increase by 1.018 times with a 1 unit increase in negative equity. When the range is between 37.804 and 62.910, the odds of default relative to continuity increase by 1.013 times with a 1 unit increase in negative equity. However, when the negative equity is above 62.910, the odds of default relative to continuity will not change with 1 unit increases.

The effect of negative equity dummy is significantly positive. The odds for mortgages with negative equity being defaulted instead of continued are 2.545 times as high as the same odds for mortgages with non-negative equity.

In most ranges, the coefficients of original loan-to-value are positive, which means the mortgages with higher loan-to-value have a higher default risk. However, when its value ranges between 0.780 and 0.800, the effect becomes insignificantly negative. When original loan-to-value is below 0.650, the coefficient is 3.226 and the standard error is 1.171. In ranges between 0.650 and 0.780, the coefficient drops to 1.725 and the standard error is 1.222. When value ranges are between 0.780 and 0.800, the coefficient becomes insignificantly negative, which is -6.195 and the standard error is 5.324. In ranges between 0.800 and 0.950, the coefficient becomes significantly positive again, which is 3.326 and the standard error is 0.749. Finally, when original loan-to-value is above 0.950, the coefficient drops to 2.375 and the standard error is 4.508.

The results of this study indicate that the credit score has a strongly positive effect on the prepayment decision and a significantly negative effect on the default decision. The coefficient is 1.784 for prepayment (the standard error is 0.670) and is -7.901 for default (the standard error is 1.469) when the credit score is below 0.688. The coefficient increases to 3.453 for prepayment (the standard error is 0.923) and to -5.916 for default (the standard error is 2.471) when the credit score ranges between 0.688 and 0.736. However, the coefficient decreases to 0.945 for prepayment (the standard error is 1.212) and increases to -2.736 for default (the standard error is 3.894) when the credit score ranges between 0.736 and 0.774. The coefficient dramatically increases to 6.870 for prepayment (the standard error is 1.637) and drops to -20.730 for default (the standard error is 6.733) when the credit score is above 0.744. These results indicate both termination risks are sensitive to the high-level credit score. The households with higher credit scores are more likely to prepay and less likely to default.

The effect of the debt-to-income ratio on the prepayment and default risk is not significant.

The dummy for the number of units has a negative effect on both termination risks. The odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.583 times as high as the same odds for mortgages covering only one house unit. Moreover, the odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.593 times as high as the same odds for mortgages covering only one house unit. The results show that the mortgages for larger houses are less likely to be prepaid or defaulted than the mortgages for smaller houses.

Log loan size has a positive effect on the prepayment risk and a negative effect on the default risk in most ranges. The odds of prepayment relative to continuity increase by 2.106 times with a 1 unit increase in log loan size when loan size is below $105,030. In this range, the odds of default relative to continuity increase by 1.405 times with a 1 unit increase in log loan size. When ranges are between $105,030 and $149,941, the odds of prepayment relative to continuity increase by 2.289 times with a 1 unit increase in log loan size. In this range, the odds of default relative to continuity decrease by 0.631 times with a 1 unit increase in log loan size. However, the odds of prepayment relative to continuity increase by 1.228 times with a 1 unit increase in log loan size when ranges are between $149,941 and $210,029. In this range, the odds of default relative to continuity decrease by 0.881 times with a 1 unit increase in log loan size. Furthermore, when size is above $210,029, the odds of prepayment relative to continuity increase by 2.662 times with a 1 unit increase in log loan size. In this range, the odds of default relative to continuity decrease by 0.968 times with a 1 unit increase in log loan size.

Generally, loan age has a strongly positive relationship with the prepayment, but its effect is insignificantly negative when its value is between 24 and 45 months. When the loan age is below 24 months, the effect of loan age on the default decision is significantly positive. However, when loan age is above 24 months, its effect on default becomes significantly negative. When loan age is below 11 months, the odds of prepayment relative to continuity increase by 1.149 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity increase by 1.517 times with a 1 unit increase in loan age. When ranges are between 11 and 24 months, the odds of prepayment relative to continuity increase by 1.002 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity increase by 1.080 times with a 1 unit increase in loan age. When loan age ranges between 24 and 45 months, the odds of prepayment relative to continuity decrease by 0.998 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity decrease by 0.994 times with a 1 unit increase in loan age. When loan age is above 45 months, the odds of prepayment relative to continuity increase again by 1.006 times with a 1 unit increase in loan age, and the odds ratio for default remains the same as 0.994 in this range.

The time period in each model is also controlled. The results show mortgages are more likely to be prepaid from March 1999 to July 2003 and are less likely to be prepaid from July 2003 to November 2008. Prepayment increases again after November 2008. On the other hand, before January 2008, mortgages are less likely to default. However, starting from January 2008, mortgages become more likely to default.

Table 50 also shows the results of the competing risks of prepayment and 90-days-delinquency based on model 2.12. The effect of each explanatory variable is almost the same as the effect in the model using prepayment and default as the dependent variables, with the absolute value of the coefficients slightly changed. However, the dummy for the number of units has a smaller effect on the 90-days-delinquency risk, the odds for mortgages covering more than one house unit being delinquent instead of continued are about 0.969 times as high as the same odds for mortgages covering only one house unit. Moreover, the 90-days-delinquency risk continues to increase starting from March 1999.

Model 2: The competing risks model based on a class of transformation survival models

Tables 51 and 52 show the monthly and yearly results estimated by model 2.20. The results of the model using monthly prepayment and default as dependent variables are first explained in detail. The difference of the results of the model using monthly prepayment and 90-days-delinquency as dependent variables is then discussed. A general comparison of monthly and yearly results is listed at the end.

[insert Table 51 here]

Table 51. Monthly results for Phoenix: the competing risks models based on a class of discrete transformation survival models for default and 90-days-delinqeuncy with maximum cd

	The competing risks of the prepayment and the default		The competing risks of the prepayment and the 90-days-delinqeuncy	
	Prepayment	Default	Prepayment	90-days-delinqeuncy
Prepayment penalty	-0.796 (0.231)	--	-0.808 (0.231)	--
Call option variables				
Call option part 1	5.357 (0.909)	--	5.378 (0.917)	--
Call option part 2	7.403 (0.883)	--	7.256 (0.901)	--
Call option part 3	7.826 (0.659)	--	8.271 (0.667)	--
Call option part 4	1.85 (0.428)	--	2.129 (0.425)	--
The unemployment rate variables				
The unemployment rate part 1	--	0.169 (0.144)	--	0.316 (0.102)
The unemployment rate part 2	--	0.142 (0.127)	--	0.338 (0.119)
Negative equity variables				
Negative equity part 1	--	0.019 (0.025)	--	0.025 (0.025)
Negative equity part 2	--	0.047 (0.020)	--	0.042 (0.019)
Negative equity part 3	--	0.034 (0.017)	--	0.052 (0.016)
Negative equity part 4	--	0.004 (0.010)	--	0.002 (0.009)
The negative equity dummy	--	1.445 (0.363)	--	1.176 (0.336)
Original loan-to-value variables				
Original LTV part 1	--	4.007 (1.716)	--	2.641 (1.289)
Original LTV part 2	--	4.541 (2.311)	--	2.643 (2.038)
Original LTV part 3	--	-26.494 (11.748)	--	-26.752 (10.352)
Original LTV part 4	--	7.701 (1.816)	--	6.664 (1.548)
Original LTV part 5	--	13.581 (16.742)	--	19.727 (14.733)
Log loan size variables				
Log loan size part 1	0.746 (0.075)	0.523 (0.563)	1.343 (0.659)	-16.949 (3.182)
Log loan size part 2	0.825 (0.121)	-1.489 (0.822)	3.482 (0.925)	-25.734 (5.483)
Log loan size part 3	0.208	0.714	0.857	-3.862

	(0.130)	(0.811)	(1.244)	(7.484)
Log loan size part 4	0.978	-0.154	6.658	-38.376
	(0.093)	(0.551)	(1.653)	(11.108)
Loan age variables				
Loan age part 1	0.139	0.411	0.777	0.267
	(0.008)	(0.170)	(0.076)	(0.426)
Loan age part 2	0.002	0.108	0.823	-0.617
	(0.003)	(0.030)	(0.122)	(0.677)
Loan age part 3	-0.002	0.005	0.216	-0.041
	(0.002)	(0.012)	(0.132)	(0.714)
Loan age part 4	0.006	-0.012	0.971	-0.349
	(0.001)	(0.006)	(0.093)	(0.489)
The credit score variables				
The credit score part 1	1.768	-17.092	0.138	0.247
	(0.664)	(3.764)	(0.008)	(0.065)
The credit score part 2	3.449	-19.683	0.003	0.061
	(0.922)	(6.000)	(0.004)	(0.024)
The credit score part 3	0.932	-2.364	-0.002	-0.002
	(1.206)	(7.861)	(0.002)	(0.011)
The credit score part 4	6.843	-39.567	0.006	-0.009
	(1.647)	(12.507)	(0.001)	(0.005)
The debt-to-income ratio variables				
DTI part 1	-0.839	0.218	-0.861	2.163
	(0.372)	(2.482)	(0.372)	(2.237)
DTI part 2	0.928	6.627	1.013	6.661
	(0.494)	(3.104)	(0.495)	(2.771)
DTI part 3	-0.671	0.495	-0.679	2.342
	(0.538)	(3.174)	(0.540)	(2.724)
DTI part 4	0.316	2.758	0.356	1.798
	(0.344)	(1.816)	(0.345)	(1.575)
The dummy for the number of units	-0.545	-1.147	-0.601	0.863
	(0.224)	(1.301)	(0.230)	(0.824)
Time period variables				
Time period part 1	0.022	-0.008	0.022	0.003
	(0.002)	(0.005)	(0.002)	(0.003)
Time period part 2	-0.037	0.080	-0.036	0.009
	(0.001)	(0.024)	(0.001)	(0.021)
Time period part 3	0.017	-0.006	0.016	0.006
	(0.001)	(0.018)	(0.001)	(0.017)
Prepayment dummy	-15.853	--	-15.905	--
	(0.962)		(0.967)	
Default dummy	--	-11.121	--	-5.876
		(7.127)		(5.163)
Maximum gd	--	5.592	--	5.452
		(0.216)		(0.190)
	Number of mortgages 12,734		Number of mortgages 12,698	

The effect of the indicator for a prepayment penalty is negative. The odds for mortgages with a prepayment penalty are only 0.451 times as high as the same odds for mortgages without a prepayment penalty.

The effect of the value of the call option is always significantly positive. When the value of the call option is below -0.025, the coefficient is 5.357 (the standard error is 0.909). When the value of the call option ranges between -0.025 and 0.037, the coefficient increases to 7.403 (the standard error is 0.883). When the value of the call option ranges between 0.037 and 0.102, the coefficient again increases to 7.826 (the standard error is 0.659). However, when the value of the call option is above 0.102, the coefficient drops to 1.850 (the standard error is 0.428).

The positive effect of the monthly unemployment rate indicates the default risk is higher in the months with the higher unemployment rate. When the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.184 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity increase by 1.153 times with a 1 percent increase in the unemployment rate.

The coefficients of negative equity show an insignificant positive effect. This result is different from that given by the competing risks model based on the multinomial logit and the proportional hazards. In those two models, the effect of negative equity splines is significant. However, the effect of negative equity dummy is significantly positive. The odds for mortgages with negative equity being defaulted instead of continued are 4.242 times as high as the same odds for mortgages with non-negative equity.

In most ranges, the coefficients of original loan-to-value partially are positive, which means the mortgages with higher loan-to-value have a higher risk of default. However, when its value ranges between 0.780 and 0.800, the effect becomes insignificantly negative. When original loan-to-value is below 0.650, the coefficient is 4.007 and the standard error is 1.716. In ranges between 0.650 and 0.780, the coefficient increases to 4.541 and the standard error is 2.311. When value ranges are between 0.780 and 0.800, the coefficient becomes insignificantly negative, which is -26.494 and the standard error is 11.748. In ranges between 0.800 and 0.950, the coefficient becomes significantly positive again, which is 7.701 and the standard error is 1.816. Finally, when original loan-to-value is above 0.950, the coefficient further increases to 13.581 and the standard error is 16.742.

The results of this study indicate that the credit score has a strongly positive effect on the prepayment decision and a significantly negative effect on the default decision. The coefficient is 1.768 for prepayment (the standard error is 0.664) and is -17.092 for default (the standard error is 3.764) when the credit score is below 0.688. The coefficient increases to 3.449 for prepayment (the standard error is 0.922) and to -19.683 for default (the standard error is 6.000) when the credit score ranges between 0.688 and 0.736. However, the coefficient decreases to 0.932 for prepayment (the standard error is 1.206) and increases to -2.364 for default (the standard error is 7.861) when the credit score ranges between 0.736 and 0.774. The coefficient dramatically increases to 6.843 for prepayment (the standard error is 1.647) and drops to -39.567 for default (the standard error is 12.507) when the credit score is above 0.744. These results indicate both termination risks are sensitive to the high-level credit score. The households with high credit scores are more likely to prepay and less likely to default.

The effect of the debt-to-income ratio on the prepayment and default risks is insignificant. The dummy for the number of units has a negative effect on both prepayment and default, and the effect is significant for prepayment. The odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.580 times as high as the same odds for mortgages covering only one house unit. Moreover, the odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.318 times as high as the same odds for mortgages covering only one house unit.

Log loan size has a positive effect on the prepayment risk and a negative effect on the default risk in most ranges. The odds of prepayment relative to continuity increase by 2.109 times with a 1 unit increase in log loan size when the size is below $105,030. In this range, the odds of default relative to continuity increase by 1.687 times with a 1 unit increase in log loan size. When ranges are between $105,030 and $149,941, the odds of prepayment relative to continuity increase by 2.281 times with a 1 unit increase in log loan size. In this range, the odds of default relative to continuity decrease by 0.226 times with a 1 unit increase in log loan size. However, the odds of prepayment relative to continuity increase by 1.231 times with a 1 unit increase in log loan size when ranges are between $149,941 and $210,029. In this range, the odds of default relative to continuity increase by 2.042 times with a 1 unit increase in log loan size. Furthermore, when size is above $210,029, the odds of prepayment relative to continuity increase by 2.659 times with a 1 unit increase in log loan size. In this range, the odds of default relative to continuity decrease by 0.857 times with a 1 unit increase in log loan size.

Generally, loan age has a strongly positive relationship with the prepayment, but its effect is insignificantly negative when its value is between 24 and 45 months. When the loan age is below 45 months, the effect of loan age on the default decision is significantly positive. However, when loan age is above 45 months, its effect on default becomes significantly negative. When loan age is below 11 months, the odds of prepayment relative to continuity increase by 1.149 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity increase by 1.508 times with a 1 unit increase in loan age. When ranges are between 11 and 24 months, the odds of prepayment relative to continuity increase by 1.002 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity increase by 1.114 times with a 1 unit increase in loan age. When loan age ranges between 24 and 45 months, the odds of prepayment relative to continuity rapidly decrease by 0.998 times with a 1 unit increase in loan age. In this range, the odds of default relative to continuity increase by 1.005 times with a 1 unit increase in loan age. When loan age is above 45, the odds of prepayment relative to continuity increase again by 1.006 times with a 1 unit increase in loan age, and the odds ratio for default decrease to 0.998 in this range.

The time period in each model is also controlled. The results show mortgages are more likely to be prepaid from March 1999 to July 2003 and are less likely to be prepaid from July 2003 to November 2008. The possibility of prepayment increases again after November 2008. On the other hand, before January 2008, mortgages are less likely to default. However, starting from January 2008, mortgages become more likely to default.

Table 51 also shows the monthly results of the competing risks of prepayment and 90-days-delinquency based on model 2.20. The effect of the most explanatory variables is very similar with that in the model using prepayment and default as the dependent variable. The coefficients of the value of the call option, unemployment rate, log loan size for prepayment slightly increase, as well as the coefficients of log loan size for 90-days-delinquency decrease considerably. The effect of the credit score in this model become insignificant and the effect of log loan size become stronger.

Table 52 shows the yearly results of the competing risks of prepayment and default (90-days-delinquency) based on model 2.20. In this model, the optimum value of c_d is around 26.44.

Compared with the results using the monthly prepayment and default as dependent variables, the effect of each explanatory variable in the yearly model keeps the same with the absolute value of the coefficients slightly increase. Compared with the results using the monthly prepayment and 90-days-delinquency as dependent variables, the effect of the most explanatory variable keeps the same. However, in the yearly model, the positive effect of the credit score on prepayment and the negative effect of the credit score on default are much stronger than the monthly results. Moreover, the absolute value of the coefficients of log loan size for the 90-days-delinquecy risk drops dramatically.

[insert Tables 52 through 53 here]

Table 53 shows the comparison of the results among the competing risks models based on multinomial logit, Sueyoshi's proportional hazard, and a class of transformation survival models. The comparison of the results shows that the coefficients estimated by the model based on Sueyoshi's proportional hazards all fall into the 95 percent confidence interval of the coefficients estimated by the model based on multinomial logit. This means that, for the single-family mortgages in Phoenix, the termination risks estimated by these two models are insignificantly distinguishable. When comparing the results estimated by the model based on a class of transformation survival models with those estimated by the model based on the multinomial logit, and the prepayment and default are used as dependent variables, the coefficients for prepayment all fall into the 95 percent confidence interval; however, most of the coefficients for default fall beyond the confidence interval. When the prepayment and 90-days-delinquency are used as dependent variables, most of the coefficients for both termination risks fall beyond the confidence interval. This means the termination risks estimated by this model will be significantly different from those estimated by the other two models.

Table 52. Yearly results for Phoenix: the competing risks models based on a class of discrete transformation survival models with maximum cp and cd

	The competing risks of the prepayment and the default		The competing risks of the prepayment and the 90-days-delinqeuncy	
	Prepayment	Default	Prepayment	90-days-delinqeuncy
Prepayment penalty	-0.825 (0.213)	--	-0.829 (0.207)	--
Call option variables				
Call option part 1	7.077 (0.402)	--	7.473 (0.828)	--
Call option part 2	8.081 (0.324)	--	8.497 (0.817)	--
Call option part 3	3.679 (0.303)	--	3.708 (0.614)	--
Call option part 4	2.903 (0.314)	--	2.881 (0.378)	--
The unemployment rate variables				
The unemployment rate part 1	--	0.417 (0.127)	--	0.564 (0.090)
The unemployment rate part 2	--	0.250 (0.078)	--	0.284 (0.071)
Negative equity variables				
Negative equity part 1	--	0.025 (0.015)	--	0.053 (0.023)
Negative equity part 2	--	0.027 (0.018)	--	0.023 (0.018)
Negative equity part 3	--	0.069 (0.017)	--	0.054 (0.016)
Negative equity part 4	--	-0.001 (0.010)	--	0.004 (0.008)
The negative equity dummy	--	1.596 (0.073)	--	1.054 (0.318)
Original loan-to-value variables				
Original LTV part 1	--	4.284 (0.270)	--	2.670 (1.174)
Original LTV part 2	--	4.222 (0.299)	--	2.477 (1.825)
Original LTV part 3	--	-28.617 (3.301)	--	-25.683 (8.854)
Original LTV part 4	--	7.816 (0.410)	--	6.463 (1.341)
Original LTV part 5	--	16.982 (4.005)	--	17.918 (8.210)
Log loan size variables				

Log loan size part 1	0.722	0.726	0.744	0.290
	(0.042)	(0.199)	(0.071)	(0.379)
Log loan size part 2	0.847	-1.560	0.833	-0.574
	(0.098)	(0.353)	(0.107)	(0.583)
Log loan size part 3	0.153	0.471	0.177	0.132
	(0.114)	(0.209)	(0.116)	(0.582)
Log loan size part 4	1.035	-0.064	1.020	-0.275
	(0.084)	(0.338)	(0.080)	(0.439)
Loan age variables				
Loan age part 1	0.935	4.331	0.980	7.633
	(0.212)	(1.346)	(0.461)	(2.278)
Loan age part 2	0.616	2.190	0.617	1.267
	(0.030)	(0.284)	(0.031)	(0.172)
Loan age part 3	0.020	0.241	0.024	0.105
	(0.016)	(0.100)	(0.016)	(0.082)
Loan age part 4	0.071	-0.103	0.070	-0.075
	(0.009)	(0.052)	(0.009)	(0.050)
The credit score variables				
The credit score part 1	1.634	-16.203	1.158	-17.044
	(0.373)	(1.070)	(0.642)	(2.722)
The credit score part 2	3.371	-22.458	3.433	-24.957
	(0.590)	(0.801)	(0.883)	(3.897)
The credit score part 3	0.928	-0.654	0.815	-3.151
	(0.892)	(1.205)	(1.194)	(4.163)
The credit score part 4	7.625	-43.609	7.290	-37.850
	(1.331)	(0.799)	(1.572)	(7.204)
The debt-to-income ratio variables				
DTI part 1	-0.861	0.284	-0.886	2.139
	(0.226)	(1.655)	(0.344)	(2.047)
DTI part 2	0.957	6.776	1.068	6.755
	(0.255)	(0.646)	(0.475)	(1.867)
DTI part 3	-0.518	0.211	-0.545	2.752
	(0.259)	(1.460)	(0.513)	(2.125)
DTI part 4	0.276	3.617	0.301	1.913
	(0.227)	(0.935)	(0.289)	(1.458)
The dummy for the number of units	-0.482	-1.180	-0.523	0.673
	(0.200)	(0.108)	(0.210)	(0.933)
Time period variables				
Time period part 1	0.110	0.001	0.105	0.117
	(0.015)	(0.047)	(0.015)	(0.038)
Time period part 2	-0.426	0.357	-0.420	-0.118
	(0.008)	(0.128)	(0.009)	(0.112)
Time period part 3	0.139	0.014	0.138	-0.019
	(0.011)	((0.135)	(0.012)	(0.126)
Prepayment dummy	-12.390	--	-12.345	--
	(0.210)		(0.461)	
Default dummy	--	-14.147	--	-10.846
		(1.346)		(2.278)

Maximum gd	--	3.275 (0.113)	3.008 (0.230)
		Number of mortgages 12,734	Number of mortgages 12,698

Table 53. Results comparison for three competing risks models

Explanatory Variable	The competing risks model based on Sueyoshi's proportional hazards model	The competing risks model based on a class of transformation survival models
Prepayment penalty	The coefficient is within the confidence interval	The coefficient is within the confidence interval
The value of call option	The coefficients are within the confidence interval	The coefficients are within the confidence interval
The unemployment rate	For default, the coefficients are within the confidence interval	For default, the coefficients are within the confidence interval
	For 90-days-delinqunecy, the coefficients are within the confidence interval	For 90-days-delinquency, the coefficient of the first spline is within the confidence interval and the coefficient of the second spline is outside the interval
The value of negative equity	For default, the coefficients are within the confidence interval	For default, the coefficients of the second and the third spline are outside the confidence interval
	For 90-days-delinqunecy, the coefficients are within the confidence interval	For 90-days-delinquency, the coefficients of the second and the third spline are outside the confidence interval
The negative equity dummy	For default, the coefficient is within the confidence interval	For default, the coefficient is outside the confidence interval
	For 90-days-delinqunecy, the coefficient is within the confidence interval	For 90-days-delinquency, the coefficient is outside the confidence interval
Original LTV	For default, the coefficients are within the confidence interval	For default, the coefficients of all splines except the first one are outside the confidence interval
	For 90-days-delinqunecy, the coefficients are within the confidence interval	For 90-days-delinquency, the coefficients of the third, the fourth and the fifth spline are outside the confidence interval
Log loan size	For prepayment, the coefficients are within the confidence interval	For prepayment, when using prepayment and default as dependent variables, the coefficients are within the confidence interval. When using prepayment and 90-days-delinquency as dependent variables, the coefficients of all splines are outside the confidence interval
	For default, the coefficients are within the confidence interval	For default, the coefficients of the second and the third spline are outside the confidence interval
	For 90-days-delinqunecy, the coefficients are within the confidence interval	For 90-days-delinqunecy, the coefficients of all splines are outside the confidence interval
Loan age	For prepayment, the coefficients are within the confidence interval	For prepayment, when using prepayment and default as dependent variables, the coefficients are within the confidence interval. When using prepayment and 90-days-delinquency as dependent variables, the coefficients of all splines are outside the confidence interval
	For default, the coefficients are within the confidence interval	For default, the coefficients are within the confidence interval
	For 90-days-delinqunecy, the coefficients are within the confidence interval	For 90-days-delinqunecy, the coefficients of all splines except the first one are outside the confidence interval
The credit score	For prepayment, the coefficients are within the confidence interval	For prepayment, when using prepayment and default as dependent variables, the coefficients are within the confidence interval. When using prepayment and 90-days-delinquency as dependent variables, the coefficients of the second and the fourth spline are outside the confidence interval
	For default, the coefficients are within the	For default, the coefficients of all splines except the

	confidence interval	third one are outside the confidence interval
	For 90-days-delinqunecy, the coefficients are within the confidence interval	For 90-days-delinqunecy, the coefficients of all splines except the third one are outside the confidence interval
Debt-to-income ratio	For prepayment, the coefficients are within the confidence interval	For prepayment, when using prepayment and default as dependent variables, the coefficients are within the confidence interval. When using prepayment and 90-days-delinquency as dependent variables, the coefficients of all splines are outside the confidence interval
	For default, the coefficients are within the confidence interval	For default, the coefficients of the second and the fourth splines are outside the confidence interval
	For 90-days-delinqunecy, the coefficients are within the confidence interval	For 90-days-delinqunecy, the coefficient of the first spline is outside the confidence interval
The dummy for the number of units	For prepayment, the coefficient is within the confidence interval	For prepayment, the coefficient is within the confidence interval
	For default, the coefficient is within the confidence interval	For default, the coefficient is within the confidence interval
	For 90-days-delinqunecy, the coefficient is within the confidence interval	For 90-days-delinqunecy, the coefficient is within the confidence interval
The confidence interval indicates the 95% confidence interval of the coefficients for the competing risks model based on the multinomial logit		

Summary

This chapter presents the detail process of constructing the competing risks model based on the proportional hazards model by using Sueyoshi's methodology. This model involves the time period between the termination point and the time point *t+1* in the analysis, all of which make the estimation more accurate. Moreover, this research constructs a new competing risks model based on a class of transformation survival models by controlling the different values of the transformation parameters, the model transfers from the proportional hazards model to different frameworks, thereby making makes the results more flexible. Moreover, by estimating the maximum likelihood function, the results show the proportional hazards framework is the best model to estimate the prepayment risk, but it is not the best model to estimate the default/90-days-delinquency risk.

These two models are used to analyze the termination risks of the single-family mortgages in Phoenix. The results of both models support most of the hypothesis made by the previous literature.

For both models, the indicator for a prepayment penalty has a negative effect on the prepayment risk. The significantly positive effect of the value of the call option supports the argument that when the value of the call option is "in the money," households have more incentive to prepay their mortgages. And the positive effect of the value of negative equity supports the argument that when the value of the put option is "in the money," households have more incentive to default/delinquent their mortgages. Moreover, the months with higher unemployment rate have a higher default/delinquency risk. And the mortgages with higher loan-to-value are more likely to be defaulted/delinquent. Generally, the debt-to-income ratio has no clear effect on the prepayment risk, but it has a positive effect on the default/delinquency risk.

Consistent with the results in previous studies, the credit score has a strong positive effect on the prepayment risk and a significantly negative effect on the default risk for both models. However, its effect is not significant for the monthly 90-days-delinquency risk when using the model based on a class of transformation survival models. In this model, log loan size has a much stronger effect on the monthly 90-days-delinquency risk compared with the results estimated by the model based on Sueyoshi's proportional hazards model.

The comparison of the results among the three models shows the coefficients estimated by the model based on Sueyoshi's proportional hazards and those estimated by the model based on the multinomial logit are insignificantly distinguishable in sign. However, the coefficients estimated by the model based on a class of transformation survival models are significantly different from those estimated by the other two models. This research continues in Chapter 3 by involving the unobserved heterogeneity into the analysis.

Chapter 3

Introduction

Unobserved heterogeneity is an important component that should be considered in the modeling process, even though it is not commonly involved in the analysis of the termination risks of the mortgages. Follain, Ondrich, and Sinha (1995) involve a multiplicative form of unobserved heterogeneity θ_i into a proportional hazards model to analyze the prepayment risk of multifamily mortgages. The results clearly show that the coefficients estimated by the model controlling for unobserved heterogeneity are significantly larger than those estimated by the model that does not control for unobserved heterogeneity. Deng, Quigley and Van Order (2000) control unobserved heterogeneity in a competing risks model to analyze the termination risks of the thirty-year, fixed-rate, single-family mortgages. They find that unobserved heterogeneity among borrowers plays an important role in accounting for the borrowers' termination behavior, particularly with respect to the prepayment behavior. Wenyi Huang and Jan Ondrich (2002) use the same model to analyze the termination risks of the multifamily mortgages. Comparing the coefficients estimated by the models with and without unobserved heterogeneity, the authors present that the absolute value of the coefficients increases sharply in the model with unobserved heterogeneity. Previous literature clearly shows that the estimated prepayment hazard and default hazard can be largely different between models with and without controls for unobserved heterogeneity.

This paper uses latent classes to control unobserved heterogeneity of two different groups of borrowers and constructs three competing risks models based on the multinomial logit, the proportional hazards model and a class of transformation survival models. The models allow the coefficients of the explanatory variables to be different between two groups of borrowers by keeping the baseline the same (the coefficients of the loan age splines are the same for the two groups of borrowers). This study uses these three models to analyze the competing risks of prepayment and default of the single-family mortgages in Phoenix and compares the average conditional hazard for prepayment, default and 90-days-delinquency estimated by models controlling for unobserved heterogeneity with those estimated by models that do not control for unobserved heterogeneity. The results show that when the loan age is between 79 and 93 months, between 103 and 118 months and above 165 months, models that do not control for unobserved heterogeneity underpredict the average conditional prepayment hazard compared with models that control for unobserved heterogeneity. Moreover, when the loan age is between 120 and 165 months, models that do not control for unobserved heterogeneity highly overpredict the prepayment hazard. For the average conditional default and 90-days-delinquency hazard, models that do not control for unobserved heterogeneity overpredict the average conditional hazard compared with models that control for unobserved heterogeneity between around 49 and 94 months. The average conditional hazard for default/90-days-delinquency is very similar between two models in all other ranges.

Because the default rate remains high during the 2007-2009 Housing Bust, another interesting question concerns the shape of the default/90-days-delinquency hazard in the absence of a boom and bust. This paper constructs a simulation that assumes that housing prices remain constant after September 2004. Suppose that the value of the property changes from V_0 to V_N. To keep the original loan-to-value ratio unchanged, the size of mortgages borrowed from the bank changes from L_0 to L_N, which leads to a change in the debt-to-income ratio. Because the remaining balance of the mortgage and the value of the property change, value of the negative equity and the negative equity dummy change as well. The average conditional default hazard and the average conditional 90-days-delinquency hazard change accordingly. The paper compares the simulated conditional hazards with those estimated based on the real trend of the housing price and shows that, in the case when the housing price changes across time, the average conditional hazard dramatically increases from around age month 11 and reaches the maximum hazard at around age month 54, and then sharply decreases until around age month 93. This dramatic change of the average conditional hazard trend disappears in the case when the housing price is assumed to be unchanged after September 2004. The simulated average conditional hazard slowly increases from age month 1 up to age month 169 with an average increase rate of 3.72 percent for default hazard and 1.80 percent for 90-days-delinquency hazard.

The average difference of the conditional hazard is approximately 0.21 percent between the two cases.

The paper is separated into five sections. Section II summarizes the previous studies on the methodology of controlling unobserved heterogeneity and develops three competing risks models based on latent classes to control unobserved heterogeneity. In section III, the dataset and results are clearly discussed. The simulation analysis is presented in section IV. The summary is offered in section V.

Literature Review

Two famous articles by Elbers and Ridder (1982) and by Heckman and Singer (1984) develop a multiplicative form of unobserved heterogeneity.

The conditional hazard with unobserved heterogeneity is assumed to be

$$\theta(t, x, \beta, v) = \Phi(x, \beta)\psi(t)v \tag{3.1}$$

in both papers, where $\Phi(x, \beta)$ is a function of observed time invariant variables x with coefficient β, $\psi(t)$ being the time dependence of the probability and v is an unobserved heterogeneity. Assume that $F(v)$is the distribution of unobserved heterogeneity v, and $Z(t) = \int_0^t \psi(u)du$, then the duration distribution corresponding to 3.1 is

$$G(t, x, \beta, v) = 1 - \int_0^\infty \exp(-\Phi(x, \beta)Z(t)v)dF(v) \tag{3.2}$$

Equation 3.2 can be determined by the dataset.

Elbers and Ridder(1982) argue that to distinguish the effect of time dependence $\psi(t)$ and the effect of unobserved heterogeneity v in the context of the proportional hazards model, the following three assumptions need to be satisfied.

Assumption 1. Unobserved heterogeneity v is non-negative with distribution function $F(v)$ and $E(v) = 1$.

Assumption 2. The function $Z(t)$ is defined on $[0, \infty)$ can be written as the integral of a non-negative integrable function $\psi(t)$ which is defined on $[0, \infty)$, i.e. $Z(t) = \int_0^t \psi(u)du$ for $t > 0$.

Assumption 3. The set $S, x \epsilon S$, is an open set in R^k. The function Φ is defined on S and is non-negative, differentiable and non-constant S.

The authors prove that, based on the above three assumptions and the differentiation of equation 3.2, $Z(t)$ and the distribution of v are uniquely identified because of the variation of individual probabilities with the explanatory variables x.

Heckman and Singer (1984) argue that the mean of unobserved heterogeneity v does not have to be finite but only requires a restriction on the tail of the true distribution. Moreover, the authors argue that the explanatory variables are not a requirement when identifying the distributions $F(v)$ and the conditional duration densities. The authors keep using the above assumption 3 in their paper and change the assumptions 1 and 2 to the following:

Assumption 1^*. Unobserved heterogeneity v is non-negative with nondefective distribution function $F(v)$. v do not possess moments of any order. However, as an absolutely continuous distributions with $E(v) = \infty$, the density function $f(v)$ has to satisfy the following tail condition

$$f(v) \sim \frac{c}{(\ln(v))^{\delta} v^{(1+\varepsilon)} L(v)} \quad 3.3$$

as $v \to \infty$, where $c > 0$, $0 < \varepsilon < 1$, $\delta \geq 0$ and $L(v)$ is slowly varying in the sense of Karamata [14]. If v is a discrete valued random variable, the relationship 3.3 needs to be replaced by the following requirement:

$$dF(v) = \begin{cases} P_k & if\ v = v_k, 0 < v_0 < v_1 < \cdots \\ 0 & otherwise\ 0 \leq k \leq \infty \end{cases} \quad 3.4$$

and $v_k \sim ck$, as $k \to \infty$, where $P_k \sim \frac{c}{(\ln(k))^{\delta} k^{(1+\varepsilon)} L(k)}$ for $0 < \varepsilon < 1$ and $\delta \geq 0$. This assumption is weaker than the assumption 1, but the tails of the true distribution are required to die off at a fast rate to make sure the equations 3.3 and 3.4 are satisfied.

Assumption 2*. A crossover condition Z is defiend as $Z \in £ = \{Z(t), t \geq 0: Z(t)$ is a nonnegative increasing function with $Z(0) = 0$ and $\exists c > 0$ and t_+ not depending on the function $Z(t)$ such that $Z(t_+) = c$ where c is a known constant. This assumption is more restricted than the previously mentioned assumption 2 and plays an important role in the identification process.

Based on the above three assumptions and the relationship between the distribution $F(v)$ and equation 3.3, the authors prove the identifiability of $F(v)$. Moreover, if the functional form of the hazard is known, such as Box-Cox hazards, by applying certain restrictions on the moments of $F(v)$, the distribution of $F(v)$ can be identified without any requirement of regressors in the model.

[14] See Feller (1971, P. 275) for a good discussion of slowly varying functions.

Meyer (1990) involves a multiplicative form of unobserved heterogeneity θ_i in an interval-censored proportional hazard model for which the log-likelihood function is

$$L(\beta,\gamma,\theta) = \sum_{i-1}^{N}\{log[\int \exp\left[-\theta\sum_{s=0}^{t-1}\exp\big(Z_i(s)\beta+\gamma(s)\big)\right]du(\theta) - \delta_i\int \exp\left[-\theta\sum_{s=0}^{t}\exp\big(Z_i(s)\beta+\gamma(s)\big)\right]du(\theta)]\} \qquad 3.5$$

When model 3.5 is used in the empirical studies, the most common assumption of the distribution of $\theta_i$is the gamma distribution with mean equal to one and variance equal to σ^2. Under this assumption, model 3.5 will become

$$L\left(\beta,\gamma,\sigma^2\right) = \sum_{i-1}^{N} log\left\{\left[1+\sigma^2\sum_{s=0}^{t-1}\exp\big(Z_i(s)\beta+\gamma(s)\big)\right]^{-\sigma^{-2}} - \delta_i\left[1+\sigma^2\sum_{s=0}^{t}\exp\big(Z_i(s)\beta+\gamma(s)\big)\right]^{-\sigma^{-2}}\right\} \qquad 3.6$$

Follain, Ondrich and Sinha (1995) implement model 3.6 to analyze the prepayment risk of the multifamily mortgages. The results show that the coefficients β estimated by the model controlling for unobserved heterogeneity are significantly larger than those estimated by the model that does not control for unobserved heterogeneity.

In recent studies, the competing risks model has been developed to analyze the termination risks of the mortgages. When it involves unobserved heterogeneity in the analysis, the most commonly used model in previous studies is

$$\begin{aligned} L(\beta_p,\gamma_p,\beta_d,\gamma_d) = \prod_{i=1}^{N}\Bigg\{&\left[1-\exp(-\exp\left(Z_i(t)\beta_p+\gamma_p(t)\right)) - c_i(t)\right]^{\delta_{p_i}}\left[1-\exp(-\exp\big(Z_i(t)\beta_d+\gamma_d(t)\big))\right.\\ &\left.- c_i(t)\right]^{\delta_{d_i}}\exp(-\exp\big(Z_i(t)\beta_p+\gamma_p(t)\\ &- \exp\big(Z_i(t)\beta_d+\gamma_d(t)\big))\prod_{s=0}^{t-1}\exp(-\exp\big(Z_i(s)\beta_p+\gamma_p(s) - \exp\big(Z_i(s)\beta_d+\gamma_d(s)\big)\big)\Bigg\} \end{aligned} \qquad 3.7$$

in which $c_i(t) = 0.5\left[1-\exp(-\exp\big(Z_i(t)\beta_p+\gamma_p(t) - \exp\big(Z_i(t)\beta_d+\gamma_d(t)\big)\big) + \exp\left(-\exp\left(Z_i(t)\beta_p+\gamma_p(t)\right)\right) + \exp(-\exp\big(Z_i(t)\beta_d+\gamma_d(t)\big)\right]$. Deng, Quigley and Van Order (2000) apply model 3.7 to analyze the competing risks of prepayment and default of thirty-year, fixed-rate single-family mortgages. They find that unobserved heterogeneity among borrowers plays an important role in accounting for the borrowers' termination behavior, particularly with respect to the prepayment behavior. Wenyi Huang and Jan Ondrich (2002) apply model 3.7 to analyze the competing risks of prepayment and default of multifamily mortgages. Comparing the coefficients estimated by the models with and without unobserved heterogeneity, the authors present that the absolute value of the coefficients increases sharply in the model with unobserved heterogeneity.

Muthen and Masyn (2005) involve a general latent variable in the discrete-time survival analysis to control unobserved heterogeneity. The model allows for individuals belonging to different subpopulations without the subpopulation membership being observed. By controlling the group-specific characteristics, the coefficient of the same explanatory variable can be different among subpopulations. This method offers more flexible results. The general latent variable framework is listed below.

Let c denote a latent categorical variable with K classes, $c_i \in \{1, \dots, K\}$, where $c_i = k$ if individual i belongs to class k. The class c is determined by individual characteristics z by multinomial logit regression

$$P(c_i = k|x_i) = \frac{\exp(\alpha_k + \gamma_k x_i)}{\sum_{s=1}^{K} \exp(\alpha_s + \gamma_s x_i)} \qquad 3.8$$

where α_k is the logit intercept and γ_k is the logit slope. The last class is a reference class with $\alpha_k = 0$ and $\gamma_k = 0$. The hazard conditional on individual i belongs to class k is $P(T_i \leq t|c_i = k, x_i)$, and the survival function conditional on individual i belongs to class k is $P(T_i > t|c_i = k, x_i)$. Then the likelihood function for the entire sample is

$$L = \prod_{i=1}^{n} P(T_i \leq t|x_i)^{\delta_I} \prod_{s=0}^{t-1} P(T_i > s|x_i) \qquad 3.9$$

This paper implements the structure of latent variables in three competing risks models—the model based on the multinomial logit, the model based on the proportional hazards model and the model based on a class of transformation survival models. To make the analysis simpler, this paper assumes the latent variables belong to two categories and the baseline (loan age splines) does not change between two categories. The models are used to analyze the competing risks of prepayment and default/90-days-delinquency of the single-family mortgages in Phoenix.

Let C_{ij} denote a latent categorical variable with two classes. The probability of individual i belongs to the class j is defined by $P_r\left(C_{ij} = 1\right) = P_j$ and the relationship between P_1 and P_2 is $P_1 + P_2 = 1$.

Let T_i denote the termination month for mortgage i. No termination month can exceed the number of months in the term of a mortgage or a censoring month. Let k_i denote the minimum of the number of months in the term of the mortgage and the censoring month. Therefore,

$$T_i = \min({T_i}^D, {T_i}^P, k_i), \qquad 3.10$$

where ${T_i}^D$ denotes the random termination month of default for mortgage i, and ${T_i}^p$denotes the random termination month of prepayment for mortgage i.

Let $Z_{pi}(t)$ and $Z_{di}(t)$ be the d-dimensional time-variant covariate vector that influences the prepayment and the default decision, respectively, for the mortgagor at time t. Moreover, these covariate vectors also determine which class individual i would belong to. α is the baseline, which does not change between the two groups of individuals. In this paper, the baseline is the loan age spline. Finally, let δ_p be the indicator for a prepayment, δ_d be the indicator for whether default and δ_c be the indicator for no prepayment or default.

The following constructs the likelihood function and the average conditional hazard with latent classes for the competing risks models based on the multinomial logit, the proportional hazards, and a class of transformation survival models.

1. The likelihood function and the average conditional hazard with latent classes for the competing risks model based on the multinomial logit

The probability of a mortgage i belonging to class j prepaying at time t conditional on surviving past time $t-1$ is given by:

$$h^P\left(t|C_{ij}=1\right)=\frac{\exp(\alpha^P+Z_{pi}(t)\beta_j{}^P)}{1+\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)+\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)},\quad t=1\ldots T_i \qquad 3.11$$

The probability of a mortgage i belonging to class j defaulting at time t conditional on surviving past time $t-1$ is given by:

$$h_i{}^D\left(t|C_{ij}=1\right)=\frac{\exp(\alpha^D+Z_{di}(t)\beta_j{}^D)}{1+\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)+\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)},\quad t=1\ldots T_i \qquad 3.12$$

The probability of a mortgage i surviving past time t and belongs to class j is given by:

$$S_i(t/C_{ij}=1)=\frac{1}{1+\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)+\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)} \qquad 3.13$$

The likelihood of a mortgage i continuing past time t is given by:

$$l_i{}^C(t)=\sum_{j=1}^{2}P_r\left(C_{ij}=1\right)\times\prod_{s=1}^{t}S_i(s/C_{ij}=1)$$

$$= \sum_{j=1}^{2} P_j \times \prod_{s=1}^{t} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s){\beta_j}^D\right) + \exp\left(\alpha^P + Z_{pi}(s){\beta_j}^P\right)}, \quad t = 1 \ldots T_i \qquad 3.14$$

The likelihood of a mortgage i prepaying at time t is:

$$l_i{}^P(t) = \sum_{j=1}^{2} P_r\left(C_{ij} = 1\right) \times \ h^P\left(t/C_{ij} = 1\ \right) \ \times \prod_{s=1}^{t-1} S_i(s/C_{ij} = 1)$$

$$= \sum_{j=1}^{2} P_j \times \frac{\exp(\alpha^P + Z_{pi}(t){\beta_j}^P)}{1 + \exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right) + \exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right)}$$

$$\times \prod_{s=1}^{t-1} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s){\beta_j}^D\right) + \exp\left(\alpha^P + Z_{pi}(s){\beta_j}^P\right)}, \quad t = 1 \ldots T_i \qquad 3.15$$

And the likelihood of a mortgage i defaulting at time t is:

$$l_i{}^D(t) = \sum_{j=1}^{2} P_r\left(C_{ij} = 1\right) \times \ h^D\left(t/C_{ij} = 1\ \right) \ \times \prod_{s=1}^{t-1} S_i(s/C_{ij} = 1)$$

$$= \sum_{j=1}^{2} P_j \times \frac{\exp(\alpha^D + Z_{di}(t){\beta_j}^D)}{1 + \exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right) + \exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right)}$$

$$\times \prod_{s=1}^{t-1} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s){\beta_j}^D\right) + \exp\left(\alpha^P + Z_{pi}(s){\beta_j}^P\right)}, \quad t = 1 \ldots T_i \qquad 3.16$$

Therefore, the likelihood function for a sample of N individuals is now given by:

$$l(\beta) = \prod_{i=1}^{N} l_i{}^P(t)^{\delta_p} l_i{}^D(t)^{\delta_d} l_i{}^C(t)^{\delta_c}$$

$$= \prod_{i=1}^{N} \sum_{j=1}^{2} P_j \times \left(\frac{\exp(\alpha^P + Z_{pi}(t)\beta_j{}^P)}{1 + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)} \right)^{\delta_p}$$

$$\times \left(\frac{\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)}{1 + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)} \right)^{\delta_d}$$

$$\times \left(\frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right)} \right)^{\delta_c}$$

$$\times \prod_{s=1}^{\mathrm{t}-1} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right)} \qquad 3.17$$

And the log-likelihood function is

$$L(\beta) = \sum_{i=1}^{N} \sum_{j=1}^{2} P_j \times \left(\frac{\exp(\alpha^P + Z_{pi}(t)\beta_j{}^P)}{1 + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)} \right)^{\delta_p}$$

$$\times \left(\frac{\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)}{1 + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)} \right)^{\delta_d}$$

$$\times \left(\frac{1}{1 + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)} \right)^{\delta_c}$$

$$\times \prod_{s=1}^{\mathrm{t}-1} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right)} \qquad 3.18$$

Without heterogeneity, the conditional hazard for prepayment is:

$$h_i{}^P(t) = \Pr(T_i < t/T_i \geq t-1), \qquad t = 1 \ldots T_i$$

$$= \frac{\exp(\alpha^P + Z_{pi}(t)\beta^P)}{1 + \exp(\alpha^D + Z_{di}(t)\beta^D) + \exp\left(\alpha^P + Z_{pi}(t)\beta^P\right)}, \qquad t = 1 \ldots T_i \qquad 3.19$$

And the conditional hazard for default is:

$$h_i{}^D(t) = \Pr(T_i < t/T_i \geq t-1)\,, \qquad t = 1 \ldots T_i$$

$$= \frac{\exp(\alpha^D + Z_{di}(t)\beta^D)}{1 + \exp(\alpha^D + Z_{di}(t)\beta^D) + \exp\left(\alpha^P + Z_{pi}(t)\beta^P\right)}, \qquad t = 1 \ldots T_i \qquad 3.20$$

With heterogeneity, the prepayment hazard conditional on individual i belonging to group j is defined by equation 3.11 and the default hazard conditional on individual i belonging to group j is defined by equation 3.12. This paper assumes that the conditional hazard for prepayment at $t = 1$ is $h_i{}^P(1) = \sum_{j=1}^{2} P_j h_{ij}{}^P(1)$ and the conditional hazard for default at $t = 1$ is $h_i{}^D(1) = \sum_{j=1}^{2} P_j h_{ij}{}^D(1)$. This assumption also can be used in the models based on proportional hazard and a class of transformation survival models.

Therefore, when $t > 1$, the conditional hazard for prepayment is given by:

$$h_i{}^P(t) = \sum_{j=1}^{2} \frac{\exp(\alpha^P + Z_{pi}(t)\beta_j{}^P)}{1 + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)} \times \frac{P_j \prod_{s=1}^{t-1} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right)}}{\sum_{k=1}^{2} P_k \prod_{s=1}^{t-1} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right)}} \qquad 3.21$$

And when $t > 1$ the conditional hazard for default is given by:

$$h_i{}^D(t) = \sum_{j=1}^{2} \frac{\exp(\alpha^D + Z_{pi}(t)\beta_j{}^D)}{1 + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)} \times \frac{P_j \prod_{s=1}^{t-1} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right)}}{\sum_{k=1}^{2} P_k \prod_{s=1}^{t-1} \frac{1}{1 + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right) + \exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right)}} \qquad 3.22$$

The average conditional hazard with latent classes for prepayment will be $\bar{h}^P(t) = \frac{1}{N}\sum_{i=1}^{N} h_i{}^P(t)$, and the average conditional hazard with latent classes for default will be $\bar{h}^D(t) = \frac{1}{N}\sum_{i=1}^{N} h_i{}^D(t)$.

2. The likelihood function and the average conditional hazard with latent classes for the competing risks model based on the proportional hazards model

The probability of a mortgage i belonging to class j prepaying at time t conditional on surviving past time $t - 1$ is given by:

$$h_i{}^P\left(t/C_{ij}=1\right)=\frac{1}{1+\frac{\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)}{\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)}}\left(1-\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)+\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)\right)\right)\right) \quad 3.23$$

The probability of a mortgage i belonging to class j defaulting at time t conditional on surviving past time $t-1$ is given by:

$$h^D\left(t/C_{ij}=1\right)=\frac{1}{1+\frac{\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)}{\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)}}\left(1-\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)+\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)\right)\right)\right) \quad 3.24$$

The probability of a mortgage i belonging to class j surviving past time t is given by:

$$S_i(t/C_{ij}=1)=\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)+\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)\right)\right) \quad 3.25$$

The likelihood of a mortgage i continuing past time t is given by:

$$l_i{}^C(t)=\sum_{j=1}^{2}P_r\left(C_{ij}=1\right)\times\prod_{s=1}^{\mathrm{t}}S_i(s/C_{ij}=1)$$

$$=\sum_{j=1}^{2}P_j\times\prod_{s=1}^{\mathrm{t}}\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(s)\beta_j{}^P\right)+\exp\left(\alpha^D+Z_{di}(s)\beta_j{}^D\right)\right)\right), t=1\ldots T_i \quad 3.26$$

The likelihood of a mortgage i prepaying at time t is given by:

$$l_i{}^P(t)=\sum_{j=1}^{2}P_r\left(C_{ij}=1\right)\times h^P\left(t/C_{ij}=1\right)\times\prod_{s=1}^{\mathrm{t-1}}S_i(s/C_{ij}=1)$$

$$=\sum_{j=1}^{2}P_j\times\frac{1}{1+\frac{\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)}{\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)}}\left(1-\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)+\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right)\right)\right)\right)$$

$$\times\prod_{s=1}^{\mathrm{t-1}}\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(s)\beta_j{}^P\right)+\exp\left(\alpha^D+Z_{di}(s)\beta_j{}^D\right)\right)\right),\quad t=1\ldots T_i \quad 3.27$$

And the likelihood of a mortgage i defaulting at time t is given by:

$$l_i^{\,D}(t) = \sum_{j=1}^{2} P_r\big(C_{ij} = 1\big) \times\ h^D\big(t/C_{ij} = 1\,\big)\ \times \prod_{s=1}^{\mathrm{t}-1} S_i(s/C_{ij} = 1)$$

$$= \sum_{j=1}^{2} P_j \times \frac{1}{1 + \dfrac{\exp\big(\alpha^P + Z_{pi}(t)\beta_j^{\,P}\big)}{\exp\big(\alpha^D + Z_{di}(t)\beta_j^{\,D}\big)}} \Big(1 - \exp\Big(-\big(\exp\big(\alpha^P + Z_{pi}(t)\beta_j^{\,P}\big) + \exp\big(\alpha^D + Z_{di}(t)\beta_j^{\,D}\big)\big)\Big)\Big)$$

$$\times \prod_{s=1}^{\mathrm{t}-1} \exp\Big(-\big(\exp\big(\alpha^P + Z_{pi}(s)\beta_j^{\,P}\big) + \exp\big(\alpha^D + Z_{di}(s)\beta_j^{\,D}\big)\big)\Big), \quad t = 1 \ldots T_i \qquad 3.28$$

Therefore, the likelihood function for a sample of N individuals is now given by:

$$l(\beta) = \prod_{i=1}^{N} l_i^{\,P}(t)^{\delta_p} l_i^{\,D}(t)^{\delta_d} l_i^{\,C}(t)^{\delta_c}$$

$$= \prod_{i=1}^{N} \sum_{j=1}^{2} P_j \times \left(\frac{1}{1 + \dfrac{\exp\big(\alpha^D + Z_{di}(t)\beta_j^{\,D}\big)}{\exp\big(\alpha^P + Z_{pi}(t)\beta_j^{\,P}\big)}} \Big(1 - \exp\Big(-\big(\exp\big(\alpha^P + Z_{pi}(t)\beta_j^{\,P}\big) + \exp\big(\alpha^D + Z_{di}(t)\beta_j^{\,D}\big)\big)\Big)\Big) \right)^{\delta_p}$$

$$\times \left(\frac{1}{1 + \dfrac{\exp\big(\alpha^P + Z_{pi}(t)\beta_j^{\,P}\big)}{\exp\big(\alpha^D + Z_{di}(t)\beta_j^{\,D}\big)}} \Big(1 - \exp\Big(-\big(\exp\big(\alpha^P + Z_{pi}(t)\beta_j^{\,P}\big) + \exp\big(\alpha^D + Z_{di}(t)\beta_j^{\,D}\big)\big)\Big)\Big) \right)^{\delta_d}$$

$$\times \Big(\exp\Big(-\big(\exp\big(\alpha^P + Z_{pi}(t)\beta_j^{\,P}\big) + \exp\big(\alpha^D + Z_{di}(t)\beta_j^{\,D}\big)\big)\Big)\Big)^{\delta_c}$$

$$\times \prod_{s=1}^{\mathrm{t}-1} \exp\Big(-\big(\exp\big(\alpha^P + Z_{pi}(s)\beta_j^{\,P}\big) + \exp\big(\alpha^D + Z_{di}(s)\beta_j^{\,D}\big)\big)\Big) \qquad 3.29$$

And the log-likelihood function is

$$L(\beta) = \sum_{i=1}^{N}\sum_{j=1}^{2} P_j \times \left(\frac{1}{1+\frac{\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)}{\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)}} \left(1-\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)+\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)\right)\right)\right) \right)^{\delta_p}$$

$$\times \left(\frac{1}{1+\frac{\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)}{\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)}} \left(1-\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)+\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)\right)\right)\right) \right)^{\delta_d}$$

$$\times \left(\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)+\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)\right)\right)\right)^{\delta_c}$$

$$\times \prod_{s=1}^{t-1} \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right)+\exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right)\right)\right) \qquad 3.30$$

Without heterogeneity, the conditional hazard for the prepayment is:

$$h_i{}^P(t) = \Pr(T_i < t/T_i \geq t-1), \qquad t = 1 \ldots T_i$$

$$= \frac{1}{1+\frac{\exp(\alpha^D + Z_{di}(t)\beta^D)}{\exp\left(\alpha^P + Z_{pi}(t)\beta^P\right)}} \left(1-\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta^P\right)+\exp(\alpha^D + Z_{di}(t)\beta^D)\right)\right)\right) \qquad 3.31$$

And the conditional hazard for the default is:

$$h_i{}^D(t) = \Pr(T_i < t/T_i \geq t-1), \qquad t = 1 \ldots T_i$$

$$= \frac{1}{1+\frac{\exp(\alpha^P + Z_{di}(t)\beta^P)}{\exp\left(\alpha^D + Z_{pi}(t)\beta^D\right)}} (1-\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta^P\right)+\exp(\alpha^D + Z_{di}(t)\beta^D)\right)\right) \qquad 3.32$$

With heterogeneity, the prepayment hazard conditional on individual i belonging to group j is defined by equation 3.23 and the default hazard conditional on individual i belonging to group j is defined by equation 3.24. Therefore, when $t > 1$, the conditional hazard for prepayment is given by:

$$h_i{}^P(t) = \sum_{j=1}^{2} \frac{1}{1 + \frac{\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)}{\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)}} (1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right) + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)\right)\right))$$

$$\times \frac{P_j \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right)\right)\right)}{\sum_{k=1}^{2} P_k \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_k{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_k{}^D\right)\right)\right)} \qquad 3.33$$

And when $t > 1$, the conditional hazard for default is given by:

$$h_i{}^D(t) = \sum_{j=1}^{2} \frac{1}{1 + \frac{\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)}{\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)}} (1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right) + \exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right)\right)\right))$$

$$\times \frac{P_j \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right)\right)\right)}{\sum_{k=1}^{2} P_k \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_k{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_k{}^D\right)\right)\right)} \qquad 3.34$$

The average conditional hazard with latent classes for prepayment will be $\bar{h}^P(t) = \frac{1}{N}\sum_{i=1}^{N} h_i{}^P(t)$, and the average conditional hazard with latent classes for default will be $\bar{h}^D(t) = \frac{1}{N}\sum_{i=1}^{N} h_i{}^D(t)$.

3. The likelihood function and the average conditional hazard with latent classes for the competing risks model based on a class of transformation survival models

The probability of a mortgage i belonging to class j prepaying at time t conditional on surviving past time $t-1$ is given by:

$$h_i{}^P(t/C_{ij} = 1) = \frac{1}{1 + \frac{\frac{1}{c_{dj}}\log(1 + c_{dj}\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right))}{\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right)}}$$

$$\times\left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_j{}^P\right) + \frac{1}{c_{dj}}\log(1 + c_{dj}\exp\left(\alpha^D + Z_{di}(t)\beta_j{}^D\right))\right)\right)\right) \qquad 3.35$$

The probability of a mortgage i belonging to class j defaulting at time t conditional on surviving past time $t-1$ is given by:

$$h^D\left(t/C_{ij}=1\right)=\frac{1}{1+\frac{\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)}{\frac{1}{c_{dj}}\log(1+c_{dj}\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right))}}$$

$$\times\left(1-\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)+\frac{1}{c_{dj}}\log(1+c_{dj}\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right))\right)\right)\right) \qquad 3.36$$

The probability of a mortgage i belonging to class j surviving past time t is given by:

$$S_i(t/C_{ij}=1)=\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t)\beta_j{}^P\right)+\frac{1}{c_{dj}}\log(1+c_{dj}\exp\left(\alpha^D+Z_{di}(t)\beta_j{}^D\right))\right)\right) \qquad 3.37$$

After estimating c_{d1} and c_{d2}, the result for c_{d2} was found to be close to zero, which makes the proportional hazards model become the best model to analyze the default risk for group 2 individuals. Therefore, for group 1 individuals, the conditional hazard for the prepayment, the conditional hazard for the default and the survival functions are shown in 3.35, 3.36 and 3.37. For group 2 individuals, the conditional hazard for the prepayment, the conditional hazard for the default, and the survival functions are shown in 3.23, 3.24 and 3.25. The likelihood of a mortgage i continuing past time t is given by:

$$l_i{}^C(t)=\sum_{j=1}^{2}P_r\left(C_{ij}=1\right)\times\prod_{s=1}^{t}S_i(s/C_{ij}=1)$$

$$=P_1\times\prod_{s=1}^{t}\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(s)\beta_1{}^P\right)+\frac{1}{c_{dj}}\log\left(1+c_{dj}\exp\left(\alpha^D+Z_{di}(s)\beta_1{}^D\right)\right)\right)\right)$$

$$+P_2\times\prod_{s=1}^{t}\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(s)\beta_2{}^P\right)+\exp\left(\alpha^D+Z_{di}(s)\beta_2{}^D\right)\right)\right),t=1\ldots T_i \qquad 3.38$$

The likelihood of a mortgage i prepaying at time t is given by:

$$l_i{}^P(t)=\sum_{j=1}^{2}P_r\left(C_{ij}=1\right)\times h^P\left(t/C_{ij}=1\right)\times\prod_{s=1}^{t-1}S_i(s/C_{ij}=1)$$

$$= P_1 \times \frac{1}{1 + \frac{\frac{1}{c_{d1}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_1}^D\right)\right)}{\exp\left(\alpha^P + Z_{pi}(t){\beta_1}^P\right)}}$$

$$\times \left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_1}^P\right) + \frac{1}{c_{d1}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_1}^D\right)\right)\right)\right)\right)$$

$$\times \prod_{s=1}^{t-1} \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(s){\beta_1}^P\right) + \frac{1}{c_{dj}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(s){\beta_1}^D\right)\right)\right)\right)$$

$$+ P_2 \times \frac{1}{1 + \frac{\exp\left(\alpha^D + Z_{di}(t){\beta_2}^D\right)}{\exp\left(\alpha^P + Z_{pi}(t){\beta_2}^P\right)}} \left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_2}^P\right) + \exp\left(\alpha^D + Z_{di}(t){\beta_2}^D\right)\right)\right)\right)$$

$$\times \prod_{s=1}^{t-1} \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(s){\beta_2}^P\right) + \exp\left(\alpha^D + Z_{di}(s){\beta_2}^D\right)\right)\right), \quad t = 1 \dots T_i \qquad 3.39$$

And the likelihood of a mortgage i defaulting at time t is given by:

$${l_i}^D(t) = \sum_{j=1}^{2} P_r\left(C_{ij} = 1\right) \times \ h^D\left(t/C_{ij} = 1\right) \ \times \prod_{s=1}^{t-1} S_i(s/C_{ij} = 1)$$

$$= P_1 \times \frac{1}{1 + \frac{\exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right)}{\frac{1}{c_{dj}} \log(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right))}}$$

$$\times \left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_1}^P\right) + \frac{1}{c_{d1}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_1}^D\right)\right)\right)\right)\right)$$

$$\times \prod_{s=1}^{t-1} \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(s){\beta_1}^P\right) + \frac{1}{c_{dj}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(s){\beta_1}^D\right)\right)\right)\right)$$

$$+ P_2 \times \frac{1}{1 + \frac{\exp\left(\alpha^P + Z_{pi}(t){\beta_2}^P\right)}{\exp\left(\alpha^D + Z_{di}(t){\beta_2}^D\right)}} \left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_2}^P\right) + \exp\left(\alpha^D + Z_{di}(t){\beta_2}^D\right)\right)\right)\right)$$

$$\times \prod_{s=1}^{t-1} \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(s){\beta_2}^P\right) + \exp\left(\alpha^D + Z_{di}(s){\beta_2}^D\right)\right)\right), \quad t = 1 \dots T_i \qquad 3.40$$

Therefore, the likelihood function for a sample of N individuals is now given by:

$$l(\beta)=\prod_{i=1}^{N} {l_i}^{P}(t)^{\delta_p}{l_i}^{D}(t)^{\delta_d}{l_i}^{C}(t)^{\delta_c}$$

$$=\prod_{i=1}^{N} P_1\times\left(\frac{1}{1+\frac{\frac{1}{c_{d1}}\log\left(1+c_{dj}\exp\left(\alpha^D+Z_{di}(t){\beta_1}^D\right)\right)}{\exp\left(\alpha^P+Z_{pi}(t){\beta_1}^P\right)}}\right.$$

$$\left.\times\left(1-\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t){\beta_1}^P\right)+\frac{1}{c_{d1}}\log\left(1+c_{dj}\exp\left(\alpha^D+Z_{di}(t){\beta_1}^D\right)\right)\right)\right)\right)\right)^{\delta_p}$$

$$\times\left(\frac{1}{1+\frac{\exp\left(\alpha^P+Z_{pi}(t){\beta_j}^P\right)}{\frac{1}{c_{dj}}\log(1+c_{dj}\exp\left(\alpha^D+Z_{di}(t){\beta_j}^D\right))}}\right.$$

$$\left.\times\left(1-\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t){\beta_1}^P\right)+\frac{1}{c_{d1}}\log\left(1+c_{dj}\exp\left(\alpha^D+Z_{di}(t){\beta_1}^D\right)\right)\right)\right)\right)\right)^{\delta_d}$$

$$\times\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t){\beta_1}^P\right)+\frac{1}{c_{dj}}\log\left(1+c_{dj}\exp\left(\alpha^D+Z_{di}(t){\beta_1}^D\right)\right)\right)\right)^{\delta_c}\times$$

$$\times\prod_{s=1}^{\mathrm{t}-1}\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(s){\beta_1}^P\right)+\frac{1}{c_{dj}}\log\left(1+c_{dj}\exp\left(\alpha^D+Z_{di}(s){\beta_1}^D\right)\right)\right)\right)+P_2$$

$$\times\left(\frac{1}{1+\frac{\exp\left(\alpha^D+Z_{di}(t){\beta_j}^D\right)}{\exp\left(\alpha^P+Z_{pi}(t){\beta_j}^P\right)}}\left(1-\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t){\beta_j}^P\right)+\exp\left(\alpha^D+Z_{di}(t){\beta_j}^D\right)\right)\right)\right)\right)^{\delta_p}$$

$$\times\left(\frac{1}{1+\frac{\exp\left(\alpha^P+Z_{pi}(t){\beta_j}^P\right)}{\exp\left(\alpha^D+Z_{di}(t){\beta_j}^D\right)}}\left(1-\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t){\beta_j}^P\right)+\exp\left(\alpha^D+Z_{di}(t){\beta_j}^D\right)\right)\right)\right)\right)^{\delta_d}$$

$$\times\left(\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(t){\beta_j}^P\right)+\exp\left(\alpha^D+Z_{di}(t){\beta_j}^D\right)\right)\right)\right)^{\delta_c}$$

$$\times\prod_{s=1}^{\mathrm{t}-1}\exp\left(-\left(\exp\left(\alpha^P+Z_{pi}(s){\beta_j}^P\right)+\exp\left(\alpha^D+Z_{di}(s){\beta_j}^D\right)\right)\right) \qquad 3.41$$

And the log-likelihood function is

$$L(\beta) = \sum_{i=1}^{N} P_1 \times \left(\frac{1}{1 + \frac{\frac{1}{c_{d1}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_1}^D\right)\right)}{\exp\left(\alpha^P + Z_{pi}(t){\beta_1}^P\right)}} \right.$$

$$\left. \times \left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_1}^P\right) + \frac{1}{c_{d1}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_1}^D\right)\right)\right)\right)\right) \right)^{\delta_p}$$

$$\times \left(\frac{1}{1 + \frac{\exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right)}{\frac{1}{c_{dj}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right)\right)}} \right.$$

$$\left. \times \left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_1}^P\right) + \frac{1}{c_{d1}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_1}^D\right)\right)\right)\right)\right) \right)^{\delta_d}$$

$$\times \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_1}^P\right) + \frac{1}{c_{dj}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(t){\beta_1}^D\right)\right)\right)\right)^{\delta_c} \times$$

$$\times \prod_{s=1}^{t-1} \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(s){\beta_1}^P\right) + \frac{1}{c_{dj}} \log\left(1 + c_{dj} \exp\left(\alpha^D + Z_{di}(s){\beta_1}^D\right)\right)\right)\right) + P_2$$

$$\times \left(\frac{1}{1 + \frac{\exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right)}{\exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right)}} \left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right) + \exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right)\right)\right)\right) \right)^{\delta_p}$$

$$\times \left(\frac{1}{1 + \frac{\exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right)}{\exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right)}} \left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right) + \exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right)\right)\right)\right) \right)^{\delta_d}$$

$$\times \left(\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t){\beta_j}^P\right) + \exp\left(\alpha^D + Z_{di}(t){\beta_j}^D\right)\right)\right)\right)^{\delta_c}$$

$$\times \prod_{s=1}^{t-1} \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(s){\beta_j}^P\right) + \exp\left(\alpha^D + Z_{di}(s){\beta_j}^D\right)\right)\right) \qquad 3.42$$

Without heterogeneity, the conditional hazard for the prepayment is:

$$h_i{}^P(t) = \Pr(T_i < t/T_i \geq t-1), \qquad t = 1 \ldots T_i$$

$$= \frac{1}{1+\frac{\frac{1}{c_d}\log(1+c_d\exp(\alpha^D + Z_{di}(t)\beta^D))}{\exp(\alpha^P + Z_{pi}(t)\beta^P)}}\left(1-\exp\left(-\left(\exp(\alpha^P + Z_{pi}(t)\beta^P) + \frac{1}{c_d}\log(1+c_d\exp(\alpha^D + Z_{di}(t)\beta^D))\right)\right)\right) \quad 3.43$$

And the conditional hazard for the default is:

$$h_i{}^D(t) = \Pr(T_i < t/T_i \geq t-1), \qquad t = 1 \ldots T_i$$

$$= \frac{1}{1+\frac{\exp(\alpha^P + Z_{pi}(t)\beta^P)}{\frac{1}{c_d}\log(1+c_d\exp(\alpha^D + Z_{di}(t)\beta^D))}}\left(1-\exp\left(-\left(\exp(\alpha^P + Z_{pi}(t)\beta^P) + \frac{1}{c_d}\log(1+c_d\exp(\alpha^D + Z_{di}(t)\beta^D))\right)\right)\right) \quad 3.44$$

With heterogeneity, the prepayment hazard conditional on individual i belonging to group 1 is defined by equation 3.35 and the default hazard conditional on individual i belonging to group 1 is defined by equation 3.36. Moreover, the prepayment hazard conditional on individual i belong to group 2 is defined by equation 3.23 and the default hazard conditional on individual i belong to group 2 is defined by equation 3.24. Therefore, when $t > 1$, the conditional hazard for prepayment is given by:

$$h_i{}^P(t) = \frac{1}{1+\frac{\frac{1}{c_{d1}}\log\left(1+c_{d1}\exp\left(\alpha^D + Z_{di}(t)\beta_1{}^D\right)\right)}{\exp\left(\alpha^P + Z_{pi}(t)\beta_1{}^P\right)}}\left(1-\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_1{}^P\right) + \frac{1}{c_{d1}}\log\left(1+c_{d1}\exp\left(\alpha^D + Z_{di}(t)\beta_1{}^D\right)\right)\right)\right)\right)$$

$$\times \frac{P_1\exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right) + \frac{1}{c_{dj}}\log(1+c_{dj}\exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right))\right)\right)_2}{P_1\exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right) + \frac{1}{c_{dj}}\log(1+c_{dj}\exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right))\right)\right) + P_2\exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_2{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_2{}^D\right)\right)\right)}$$

$$+ \frac{1}{1+\frac{\exp\left(\alpha^D + Z_{di}(t)\beta_2{}^D\right)}{\exp\left(\alpha^P + Z_{pi}(t)\beta_2{}^P\right)}}\left(1-\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_2{}^P\right) + \exp\left(\alpha^D + Z_{di}(t)\beta_2{}^D\right)\right)\right)\right)$$

$$\times \frac{P_2\exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_2{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_2{}^D\right)\right)\right)}{P_1\exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right) + \frac{1}{c_{dj}}\log\left(1+c_{dj}\exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right)\right)\right)\right) + P_2\exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_2{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_2{}^D\right)\right)\right)} \quad 3.45$$

And when $t > 1$, the conditional hazard for default is given by:

$$h_i{}^D(t) = \frac{1}{1+\frac{\exp\left(\alpha^P + Z_{pi}(t)\beta_1{}^P\right)}{\frac{1}{c_{d1}}\log(1+c_{d1}\exp\left(\alpha^D + Z_{di}(t)\beta_1{}^D\right))}}\left(1-\exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_1{}^P\right) + \frac{1}{c_{d1}}\log(1+c_{d1}\exp\left(\alpha^D + Z_{di}(t)\beta_1{}^D\right))\right)\right)\right)$$

$$\times \frac{P_1 \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right) + \frac{1}{c_{dj}}\log(1 + c_{dj}\exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right))\right)\right)_2}{P_1 \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right) + \frac{1}{c_{dj}}\log(1 + c_{dj}\exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right))\right)\right) + P_2 \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_2{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_2{}^D\right)\right)\right)}$$

$$+ \frac{1}{1 + \frac{\exp\left(\alpha^P + Z_{pi}(t)\beta_2{}^P\right)}{\exp\left(\alpha^D + Z_{di}(t)\beta_2{}^D\right)}}\left(1 - \exp\left(-\left(\exp\left(\alpha^P + Z_{pi}(t)\beta_2{}^P\right) + \exp\left(\alpha^D + Z_{di}(t)\beta_2{}^D\right)\right)\right)\right)$$

$$\times \frac{P_2 \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_2{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_2{}^D\right)\right)\right)}{P_1 \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_j{}^P\right) + \frac{1}{c_{dj}}\log(1 + c_{dj}\exp\left(\alpha^D + Z_{di}(s)\beta_j{}^D\right))\right)\right) + P_2 \exp\left(-\sum_{s=1}^{t-1}\left(\exp\left(\alpha^P + Z_{pi}(s)\beta_2{}^P\right) + \exp\left(\alpha^D + Z_{di}(s)\beta_2{}^D\right)\right)\right)} \quad 3.46$$

The average conditional hazard with latent classes for prepayment will be $\bar{h}^P(t) = \frac{1}{N}\sum_{i=1}^{N} h_i{}^P(t)$, and the average conditional hazard with latent classes for default will be $\bar{h}^D(t) = \frac{1}{N}\sum_{i=1}^{N} h_i{}^D(t)$.

Results

This section presents the results of using the above three competing risks models with heterogeneity to analyze the termination risks of the single-family mortgages in Phoenix. The average conditional hazards estimated by the above models are compared with those estimated by models that do not control for unobserved heterogeneity.

For each model, two sets of dependent variables are analyzed, the prepayment and the default being the first, and the prepayment and the 90-days-delinquency being the second. For both groups of dependent variables, the indicator for a prepayment penalty and the value of the call option are used to explain the prepayment risk; the unemployment rate, negative equity, a dummy for negative equity and original loan-to-value are used to explain the default/90-days-delinquency risk; the credit score, the debt-to-income ratio, log mortgage size, current loan age in months, a dummy for units and time period dummies are used to explain both risks. Moreover, the dataset used in the paper is the dataset with mortgages being right censored one month before the modification.

Results 1: The results of the competing risks model with latent classes based on the multinomial logit

Table 54 shows the results of the competing risks of prepayment and default based on model 3.18. The result shows the probability of a mortgage belonging to the group 1 individuals is 0.762 and belonging to the group 2 individuals is 0.238. Compared the important coefficients across groups, the major differences are summarized in the following. The indicator for a prepayment penalty has a negative on the prepayment risk for both groups and its effect on the group 1 mortgages is stronger. The value of call option has a positive effect on the prepayment risk and it has a lager effect on the group 2 mortgages. The negative equity dummy has a positive effect on the default/90-days-delinquency risk and its effect on the group 1 mortgages is larger. Moreover, the credit score has a positive effect on the prepayment risk and its effect on the group 2 mortgages is stronger. The detail information of the results is listed below.

[insert Table 54 here]

For groups 1 and 2 mortgages, the effect of the indicator for a prepayment penalty is negative. The odds for mortgages with the prepayment penalty are only 0.372 times as high as the same odds for mortgages without a prepayment penalty in group 1. While the odds for mortgages with the prepayment penalty are 0.732 times as high as the same odds for mortgages without a prepayment penalty in group 2. This means that the prepayment penalty has a greater impact on the group 1 mortgages.

The effect of the value of the call option is all positive and the significantly positive effect supports the argument that when the value of the call option is “in the money,” households have more incentive to prepay their mortgages. Moreover, the coefficients for the group 2 mortgages are larger than those for the group 1 mortgages.

Table 54. The monthly results for Phoenix based on the multinomial logit with prepayment and default as dependent variables (Baseline is loan age, P_1 =0.762, P_2 = 0.238)

	Group 1 Mortgages		Group 2 Mortgages	
	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.988	--	-0.312	--
	(0.290)		(0.631)	
Call option variables				
Call option part 1	5.677	--	27.273	--
	(0.877)		(4.756)	
Call option part 2	7.024	--	70.861	--
	(0.978)		(15.931)	
Call option part 3	4.461	--	33.962	--
	(0.872)		(3.750)	
Call option part 4	0.200	--	0.784	--
	(0.596)		(1.074)	
The unemployment rate variables				
The unemployment rate part 1	--	0.196	--	0.051
		(0.098)		(0.055)
The unemployment rate part 2	--	0.069	--	0.118
		(0.052)		(0.006)
Negative equity variables				
Negative equity part 1	--	0.019	--	-0.041
		(0.014)		(0.014)
Negative equity part 2	--	0.022	--	-0.020
		(0.009)		(0.012)
Negative equity part 3	--	0.014	--	-0.025
		(0.005)		(0.011)
Negative equity part 4	--	0.002	--	-0.028
		(0.003)		(0.009)
The negative equity dummy	--	1.038	--	0.775
		(0.204)		(0.004)
Original loan-to-value variables				
Original LTV part 1	--	3.206	--	3.843
		(1.003)		(0.013)
Original LTV part 2	--	2.038	--	1.444
		(0.844)		(0.004)
Original LTV part 3	--	-4.365	--	-7.508
		(0.377)		(0.004)
Original LTV part 4	--	3.611	--	2.804
		(0.751)		(0.004)
Original LTV part 5	--	-1.746	--	2.904
		(1.150)		(0.004)
Log loan size variables				
Log loan size part 1	0.516	0.461	1.864	-0.303
	(0.104)	(0.251)	(0.363)	(0.223)

Log loan size part 2	0.663	-0.529	1.924	-0.562
	(0.158)	(0.419)	(0.450)	(0.006)
Log loan size part 3	-0.184	-0.400	1.614	-0.115
	(0.171)	(0.401)	(0.467)	(0.004)
Log loan size part 4	1.176	0.314	0.533	-0.037
	(0.129)	(0.274)	(0.319)	(0.004)
Loan age variables				
Loan age part 1	0.127	0.413	0.127	0.413
	(0.007)	(0.124)	(0.007)	(0.124)
Loan age part 2	0.010	0.074	0.010	0.074
	(0.003)	(0.016)	(0.003)	(0.016)
Loan age part 3	0.001	0.003	0.001	0.003
	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	0.005	0.002	0.005	0.002
	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables				
The credit score part 1	0.572	-9.275	7.866	-6.664
	(0.711)	(1.478)	(1.977)	(0.014)
The credit score part 2	1.311	-10.379	12.305	-7.110
	(0.958)	(1.679)	(1.728)	(0.004)
The credit score part 3	1.029	-1.358	0.669	-2.262
	(0.825)	(1.993)	(2.205)	(0.004)
The credit score part 4	3.964	-26.265	19.578	-24.822
	(1.108)	(3.070)	(4.003)	(0.004)
The debt-to-income ratio variables				
DTI part 1	-1.424	0.613	1.795	0.950
	(0.449)	(0.898)	(1.196)	(0.006)
DTI part 2	1.497	3.981	-0.462	3.332
	(0.563)	(0.797)	(0.982)	(0.004)
DTI part 3	-0.553	1.011	-0.791	0.619
	(0.623)	(0.898)	(0.901)	(0.004)
DTI part 4	0.369	1.452	0.162	0.778
	(0.431)	(0.816)	(0.810)	(0.004)
The dummy for the number of units	-0.298	-0.667	-1.499	-0.443
	(0.275)	(0.727)	(0.654)	(0.004)
Time period variables				
Time period part 1	0.021	-0.004	0.014	-0.951
	(0.002)	(0.003)	(0.008)	(0.012)
Time period part 2	-0.024	0.048	-0.140	-0.098
	(0.001)	(0.009)	(0.008)	(0.035)
Time period part 3	-0.001	0.008	0.139	-0.037
	(0.002)	(0.008)	(0.010)	(0.012)
Prepayment dummy	-12.025	--	-37.117	--
	(1.268)		(4.470)	
Default dummy	--	-14.767	--	-12.597
		(2.703)		(0.020)

Number of mortgages 12,734

The effect of the monthly unemployment rate for both groups is positive. For the group 1 mortgages, when the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.217 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity increase by 1.071 times with a 1 percent increase in the unemployment rate. For the group 2 mortgages, when the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.052 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity increase by 1.125 times with a 1 percent increase in the unemployment rate.

The effect of negative equity on the default risk is consistently positive for the group 1 mortgages; however, the effect is insignificant for the group 2 mortgages. Therefore, the negative equity splines are a strong factor in determining the default risk for the group 1 mortgages, but not a strong factor for the group 2 mortgages. Moreover, the effect of negative equity dummy for both groups is significantly positive, which supports that when the put option is "in the money," households have more incentive to default their mortgages. For the group 1 mortgages, the odds for mortgages with negative equity being defaulted instead of continued are 2.834 times as high as the same odds for mortgages with non-negative equity. For group 2 mortgages, the odds for mortgages with negative equity being defaulted instead of continued are 2.171 times as high as the same odds for mortgages with non-negative equity

The results of original loan-to-value partially support the previous literature that the original loan-to-value has a positive effect on the default risk. When its value ranges between 0.780 and 0.800, the effect becomes insignificant for both groups, and when its value ranges above 0.950, the effect is negative for group 1 mortgages.

Consistent with the results in previous studies, the credit score has a strong negative effect on the default risk for both groups of mortgages. Moreover, the results of this study indicate it has a strongly positive effect on the prepayment decision for both groups of mortgages. The results also indicate that both termination risks are sensitive to the high-level credit score. The households in both groups with high credit scores are more likely to prepay and less likely to default.

The effect of the debt-to-income ratio on the default risk is significantly positive for group 2 mortgages, which supports the hypothesis that household with high debt-to-income ratio are more likely to default. However, its effect on the default risk is insignificant for group 1 mortgages. Moreover, for the prepayment risk, the effect of the debt-to-income ratio is inconsistent for both groups.

The dummy for the number of units has a negative effect on both the prepayment and default risks for both groups of mortgages. For group 1, the odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.742 times as high as the same odds for mortgages covering only one house unit. Moreover, the odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.513 times as high as the same odds for mortgages covering only one house unit. For group 2, the odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.223 times as high as the same odds for mortgages covering only one house unit. Moreover, the odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.642 times as high as the same odds for mortgages covering only one house unit. The results show the mortgages for larger houses are less likely to be prepaid or defaulted than the mortgages for smaller houses. Moreover, the housing unit has a larger effect on the prepayment risk for the group 2 mortgages and has a larger effect on the default risks for the group 1 mortgages.

Log loan size has a significantly positive effect on the prepayment. However, the effect is negative when the loan size ranges between $149,941 and $210,029 for the group 1 mortgages. On the other hand, its effect on the default is negative. However, when the size is below $105,030 and above $210,029, the effect becomes positive for group 1 mortgages.

As the baseline, the loan age has the same effect for both groups of mortgages. The results in this paper show that it has a strongly positive relationship with both the prepayment and the default risks and the effect for both risks is the strongest when loan age is below 11 months.

The paper also controls for the time period. The results show that mortgages are more likely to be prepaid from March 1999 to July 2003 for both groups and are less likely to be prepaid from July 2003 to November 2008. For group 1, prepayment continues decreasing after November 2008, but for group 2, prepayment increases after November 2008. On the other hand, for group 1, before January 2008, mortgages are less likely to default. However, starting from January 2008, mortgages become more likely to default. For group 2, the likelihood of default is consistently low.

Table 55 shows the results of the competing risks of prepayment and 90-days-delinquency based on model 3.18. The effect of each explanatory variable is almost the same as the effect in the model using prepayment and default as the dependent variable, with the absolute value of the coefficients slightly changed. However, the dummy for the number of units has a positive effect on the 90-days-delinquency risk for the group 1 mortgages, the odds for mortgages covering more than one unit being delinquent instead of continued are about 1.078 times as high as the same odds for mortgages covering only one house unit.

[insert Table 55 here]

Figures 28 and 29 show the comparison of the average conditional hazard for prepayment with and without controls for unobserved heterogeneity. The average conditional hazards for prepayment in figure 28 are based on the model using prepayment and default as dependent variables, and those in figure 29 are based on the model using prepayment and 90-days-delinquency as dependent variables.

[insert Figures 28 through 31 here]

The results in figure 28 show that when the loan age is below 107 months, the average conditional hazard for the prepayment is similar for the models with and without controls for unobserved heterogeneity. When the loan age is between 107 and 119 months and above 165 months, the model that does not control for unobserved heterogeneity underpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity. Moreover, when the loan age is between 119 and 165 months, the model that does not control for unobserved heterogeneity is overpredicts the hazard compared with the model that controls for unobserved heterogeneity.

The results in figure 29 show that when the loan age is below 77 months, the average conditional hazard for the prepayment is similar across the models with and without controls for unobserved heterogeneity. When the loan age is between 77 and 95 months, between 103 and 121 months and above 165 months, the model that does not control for unobserved heterogeneity underpredicts the average conditional hazard compared with the model that controls for

Table 55. The monthly results for Phoenix based on the multinomial logit with prepayment and 90-days-delinquency as dependent variables (Baseline is loan age, $P_1 = 0.783, P_2 = 0.217$)

	Group 1 Mortgages		Group 2 Mortgages	
	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.973	--	-0.326	--
	(0.294)		(0.577)	
Call option variables				
Call option part 1	5.642	--	33.250	--
	(0.847)		(5.124)	
Call option part 2	7.212	--	81.848	--
	(0.915)		(26.159)	
Call option part 3	4.900	--	38.360	--
	(0.792)		(1.930)	
Call option part 4	0.662	--	0.968	--
	(0.571)		(0.800)	
The unemployment rate variables				
The unemployment rate part 1	--	0.223	--	0.036
		(0.071)		(0.018)
The unemployment rate part 2	--	0.153	--	0.129
		(0.047)		(0.004)
Negative equity variables				
Negative equity part 1	--	0.020	--	-0.038
		(0.013)		(0.004)
Negative equity part 2	--	0.024	--	-0.022
		(0.008)		(0.004)
Negative equity part 3	--	0.020	--	-0.023
		(0.005)		(0.004)
Negative equity part 4	--	0.000	--	-0.031
		(0.003)		(0.004)
The negative equity dummy	--	0.873	--	0.705
		(0.195)		(0.004)
Original loan-to-value variables				
Original LTV part 1	--	2.151	--	4.225
		(0.812)		(0.005)
Original LTV part 2	--	0.680	--	1.312
		(0.869)		(0.004)
Original LTV part 3	--	-2.387	--	-8.259
		(3.311)		(0.004)
Original LTV part 4	--	3.043	--	2.549
		(0.667)		(0.004)
Original LTV part 5	--	-0.824	--	3.194
		(0.383)		(0.004)
Log loan size variables				
Log loan size part 1	0.561	0.190	1.955	-0.309
	(0.086)	(0.280)	(0.257)	(0.053)

Log loan size part 2	0.661	0.028	2.036	-0.619
	(0.145)	(0.323)	(0.421)	(0.005)
Log loan size part 3	-0.100	-0.849	1.540	-0.105
	(0.157)	(0.348)	(0.446)	(0.004)
Log loan size part 4	1.132	0.234	0.552	-0.041
	(0.114)	(0.243)	(0.317)	(0.004)
Loan age variables				
Loan age part 1	0.127	0.234	0.127	0.234
	(0.007)	(0.045)	(0.007)	(0.045)
Loan age part 2	0.010	0.037	0.010	0.037
	(0.004)	(0.012)	(0.004)	(0.012)
Loan age part 3	0.001	0.003	0.001	0.003
	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	0.005	0.002	0.005	0.002
	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables				
The credit score part 1	0.415	-10.251	6.492	-6.060
	(0.699)	(1.358)	(1.623)	(0.005)
The credit score part 2	1.507	-11.562	12.428	-7.821
	(0.822)	(2.111)	(1.666)	(0.004)
The credit score part 3	0.938	-6.055	1.015	-2.056
	(0.897)	(2.323)	(2.160)	(0.004)
The credit score part 4	3.792	-24.243	19.638	-27.304
	(0.889)	(2.382)	(2.030)	(0.004)
The debt-to-income ratio variables				
DTI part 1	-1.443	2.046	2.114	0.863
	(0.422)	(1.377)	(0.954)	(0.005)
DTI part 2	1.471	4.195	0.058	3.665
	(0.548)	(1.252)	(0.671)	(0.004)
DTI part 3	-0.613	1.638	-0.757	0.563
	(0.601)	(1.237)	(1.086)	(0.004)
DTI part 4	0.492	1.253	-0.258	0.856
	(0.398)	(0.745)	(0.785)	(0.004)
The dummy for the number of units	-0.333	0.075	-1.594	-0.402
	(0.282)	(0.630)	(0.671)	(0.004)
Time period variables				
Time period part 1	0.021	0.001	0.014	-1.053
	(0.002)	(0.002)	(0.007)	(0.012)
Time period part 2	-0.025	0.015	-0.145	-0.089
	(0.001)	(0.008)	(0.006)	(0.004)
Time period part 3	0.001	0.012	0.144	-0.041
	(0.002)	(0.007)	(0.009)	(0.004)
Prepayment dummy	-12.441	--	-37.982	--
	(0.995)		(3.298)	
90-days-delinquecy dummy	--	-8.095	--	-11.455
		(3.346)		(0.007)

Number of mortgages 12,734

Figure 28. The average conditional hazard for prepayment for the competing risks model based on the multinomial logit (prepayment and default as dependent variables)

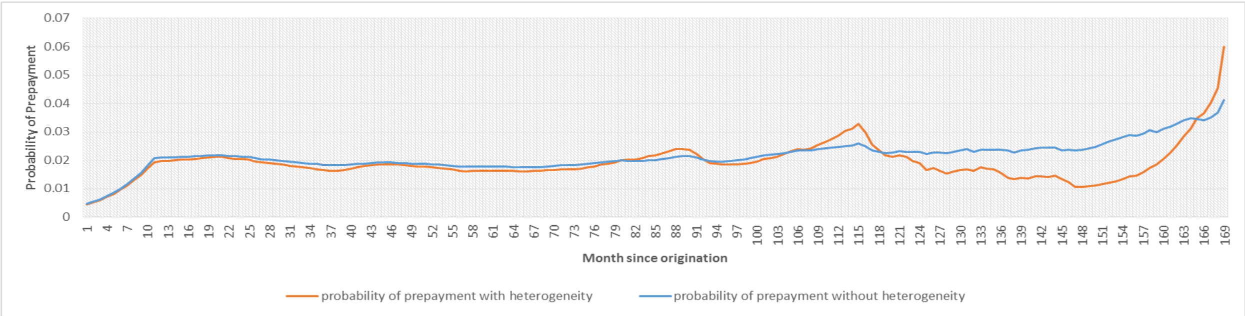


Figure 29. The average conditional hazard for prepayment for the competing risks model based on the multinomial logit (prepayment and 90-days-delinquency as dependent variables)

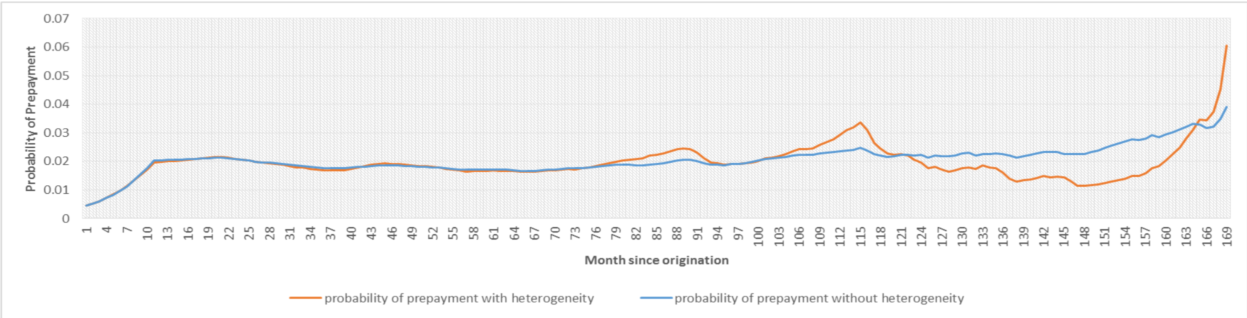


Figure 30. The average conditional hazard for default for the competing risks model based on the multinomial logit (prepayment and default as dependent variables)

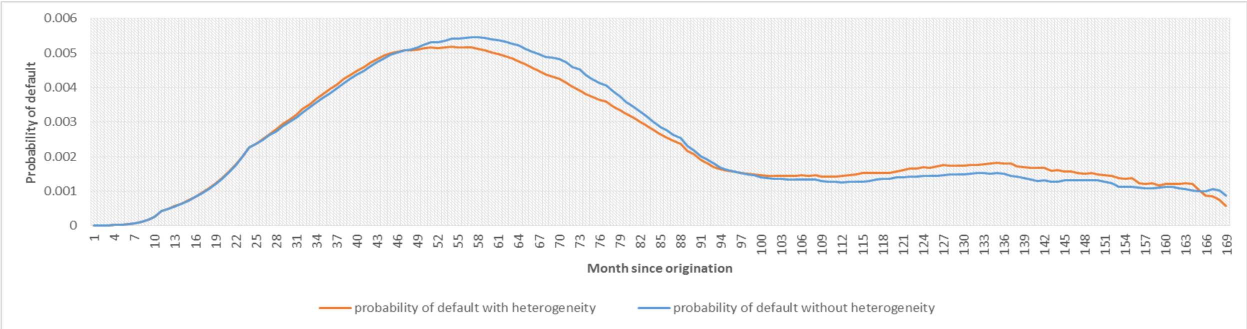
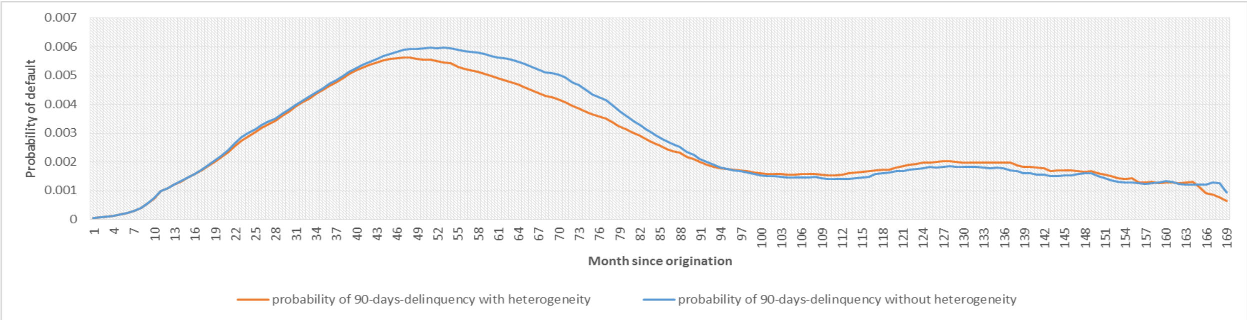


Figure 31. The average conditional hazard for 90-days-delinquency for the competing risks model based on the multinomial logit (prepayment and 90-days-delinquency as dependent variables)



unobserved heterogeneity. Moreover, when the loan age is between 121 and 165 months, the model that does not control for unobserved heterogeneity is overpredicts the hazard compared with the model that controls for unobserved heterogeneity.

Figure 30 shows the comparison of the average conditional hazard for default with and without unobserved heterogeneity. It shows that, when the loan age is below 48 months, the average conditional hazard for default is similar across the models with and without controls for unobserved heterogeneity. When the loan age is between 48 and 96 months, the model that does not control for unobserved heterogeneity overpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity. When the loan age is above 99 months, the model that does not control for unobserved heterogeneity slightly underpredicts the average conditional hazard.

Figure 31 compares the average conditional hazard for 90-days-delinquency with and without unobserved heterogeneity. It shows that, when the loan age is below 42 months and above 94 months, the average conditional hazard for 90-days-delinquency is similar across the models with and without controls for unobserved heterogeneity. When the loan age is between 42 and 94 months, the model that does not control for unobserved heterogeneity overpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity.

Results 2: The results of the competing risks model with latent classes based on the proportional hazards model

Table 56 shows the results of the competing risks of prepayment and default based on model 3.30. The results estimated by this model are very similar to the results given by model 3.18. The probability of a mortgage borrowed by the group 1 individuals is 0.763 and borrowed by the group 2 individuals is 0.237. Compared the important coefficients across groups, the major differences are summarized in the following. The indicator for a prepayment penalty has a negative on the prepayment risk for both groups and its effect on the group 1 mortgages is stronger. The value of call option has a positive effect on the prepayment risk and it has a lager effect on the group 2 mortgages. The negative equity dummy has a positive effect on the default/90-days-delinquency risk and its effect on the group 1 mortgages is larger. Moreover, the credit score has a positive effect on the prepayment risk and its effect on the group 2 mortgages is stronger. The detail information of the results is listed below.

[insert Table 56 here]

For groups 1 and 2 individuals, the effect of the indicator for a prepayment penalty is negative. The odds for mortgages with a prepayment penalty are only 0.376 times as high as the same odds for mortgages without a prepayment penalty in group 1. While the odds for mortgages with a prepayment penalty are 0.744 times as high as the same odds for mortgages without a prepayment penalty in group 2. This means the prepayment penalty has a greater impact on the group 1 mortgages.

The effect of the value of the call option is all positive and the significantly positive effect supports the argument that when the value of the call option is “in the money,” households have more incentive to prepay their mortgages. Moreover, the coefficients for the group 2 mortgages are larger than those for the group 1 mortgages.

The effect of the monthly unemployment rate for both groups is positive. For the group 1 mortgages, when the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.218 times with a 1 percent increase in the unemployment rate. When the rate is

Table56. The monthly results for Phoenix based on Sueyoshi's proportional hazards model with prepayment and default as dependent variables (Baseline is loan age, P_1 =0.763, P_2 = 0.237)

	Group 1 Mortgages		Group 2 Mortgages	
	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.978 (0.280)	--	-0.296 (0.110)	--
Call option variables				
Call option part 1	5.704 (0.509)	--	27.354 (0.129)	--
Call option part 2	6.802 (0.489)	--	70.891 (0.047)	--
Call option part 3	4.513 (0.537)	--	31.577 (3.231)	--
Call option part 4	0.276 (0.384)	--	0.677 (0.493)	--
The unemployment rate variables				
The unemployment rate part 1	--	0.197 (0.100)	--	0.051 (0.005)
The unemployment rate part 2	--	0.070 (0.055)	--	0.118 (0.005)
Negative equity variables				
Negative equity part 1	--	0.019 (0.014)	--	-0.041 (0.005)
Negative equity part 2	--	0.022 (0.008)	--	-0.020 (0.005)
Negative equity part 3	--	0.014 (0.005)	--	-0.025 (0.005)
Negative equity part 4	--	0.002 (0.003)	--	-0.028 (0.005)
The negative equity dummy	--	1.034 (0.209)	--	0.775 (0.005)
Original loan-to-value variables				
Original LTV part 1	--	3.079 (0.559)	--	3.843 (0.005)
Original LTV part 2	--	2.127 (0.318)	--	1.444 (0.005)
Original LTV part 3	--	-4.335 (0.082)	--	-7.508 (0.005)
Original LTV part 4	--	3.518 (0.442)	--	2.804 (0.005)
Original LTV part 5	--	-1.627 (0.170)	--	2.904 (0.005)
Log loan size variables				
Log loan size part 1	0.506 (0.020)	0.433 (0.114)	1.822 (0.077)	-0.303 (0.005)

Log loan size part 2	0.655	-0.512	1.822	-0.562
	(0.130)	(0.233)	(0.270)	(0.005)
Log loan size part 3	-0.183	-0.395	1.543	-0.115
	(0.156)	(0.204)	(0.191)	(0.005)
Log loan size part 4	1.161	0.300	0.520	-0.037
	(0.116)	(0.256)	(0.257)	(0.005)
Loan age variables				
Loan age part 1	0.126	0.411	0.126	0.411
	(0.007)	(0.107)	(0.007)	(0.107)
Loan age part 2	0.010	0.074	0.010	0.074
	(0.004)	(0.015)	(0.004)	(0.015)
Loan age part 3	0.001	0.003	0.001	0.003
	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	0.005	0.001	0.005	0.001
	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables				
The credit score part 1	0.534	-9.214	7.916	-6.664
	(0.227)	(0.549)	(0.344)	(0.005)
The credit score part 2	1.490	-9.948	11.322	-7.110
	(0.353)	(0.578)	(1.349)	(0.005)
The credit score part 3	0.939	-1.688	0.976	-2.262
	(0.161)	(0.630)	(0.895)	(0.005)
The credit score part 4	3.883	-25.992	19.230	-24.822
	(0.173)	(0.388)	(0.481)	(0.005)
The debt-to-income ratio variables				
DTI part 1	-1.406	0.621	1.832	0.950
	(0.341)	(0.149)	(0.479)	(0.005)
DTI part 2	1.486	3.932	-0.580	3.332
	(0.305)	(0.208)	(0.600)	(0.005)
DTI part 3	-0.554	1.015	-0.738	0.619
	(0.248)	(0.264)	(0.209)	(0.005)
DTI part 4	0.363	1.400	0.095	0.778
	(0.324)	(0.286)	(0.469)	(0.005)
The dummy for the number of units	-0.298	-0.661	-1.457	-0.443
	(0.262)	(0.128)	(0.233)	(0.005)
Time period variables				
Time period part 1	0.021	-0.004	0.013	-0.951
	(0.002)	(0.003)	(0.006)	(0.005)
Time period part 2	-0.024	0.047	-0.135	-0.098
	(0.001)	(0.009)	(0.007)	(0.005)
Time period part 3	-0.001	0.008	0.135	-0.037
	(0.002)	(0.008)	(0.009)	(0.005)
Prepayment dummy	-11.876	--	-36.541	--
	(0.180)		(0.851)	
Default dummy	--	-14.430	--	-12.597
		(0.488)		(0.005)

Number of mortgages 12,734

above 5.700, the odds of default relative to continuity increase by 1.073 times with a 1 percent increase in the unemployment rate. For the group 2 mortgages, when the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.052 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity increase by 1.125 times with a 1 percent increase in the unemployment rate.

The effect of negative equity on the default risk is consistently positive for the group 1 mortgages; however, the effect is insignificant for the group 2 mortgages. Therefore, the negative equity splines are not a strong factor for the group 2 mortgages. Moreover, the effect of negative equity dummy for both groups is significantly positive, which supports that when the put option is "in the money," households have more incentive to default their mortgages. For the group 1 mortgages, the odds for mortgages with negative equity being defaulted instead of continued are 2.812 times as high as the same odds for mortgages with non-negative equity. For group 2 mortgages, the odds for mortgages with negative equity being defaulted instead of continued are 2.171 times as high as the same odds for mortgages with non-negative equity

The results of original loan-to-value partially support the previous literature that the original loan-to-value has a positive effect on the default risk. When its value ranges between 0.780 and 0.800, the effect becomes insignificantly negative for both groups, and when its value ranges above 0.950, the effect is negative for group 1 mortgages.

Consistent with the results in previous studies, the credit score has a strong negative effect on the default risk for both groups of mortgages. Moreover, the results indicate it has a strongly positive effect on the prepayment decision for both groups of mortgages. The results also indicate that both termination risks are sensitive to the high-level credit score, which means the households with high credit scores are more likely to prepay and less likely to default.

The effect of the debt-to-income ratio on the default risk is significantly positive for both groups of mortgages, which supports the hypothesis that households with high debt-to-income ratio are more likely to default. However, for the prepayment risk, the effect of the debt-to-income ratio is inconsistent for both groups.

The dummy for the number of units has a negative effect on both the prepayment and default risks for both groups of mortgages. For group 1, the odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.742 times as high as the same odds for mortgages covering only one house unit. Moreover, the odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.516 times as high as the same odds for mortgages covering only one house unit. For group 2, the odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.233 times as high as the same odds for mortgages covering only one house unit. Moreover, the odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.642 times as high as the same odds for mortgages covering only one house unit. The results show the mortgages for larger houses are less likely to be prepaid or defaulted than the mortgages for smaller houses.

Log loan size has a significantly positive effect on the prepayment. However, the effect is negative when the loan size ranges between $149,941 and $210,029 for the group 1 mortgages. On the other hand, its effect on the default is negative. However, when the size is below $105,030 and above $210,029, the effect becomes positive for group 1 mortgages.

As the baseline, the loan age has the same effect for both groups of mortgages. The results in this paper show it has a strongly positive relationship with both the prepayment risk and the default risk and the effect for both risks is the strongest when loan age is below 11 months.

The paper also controls for the time period. The results show that mortgages are more likely to be prepaid from March 1999 to July 2003 for both groups and are less likely to be prepaid from July 2003 to November 2008. For group 1, prepayment continues decreasing after November 2008, but for group 2, prepayment increases after November 2008. On the other hand before January 2008 for group 1, mortgages are less likely to default. However, starting from January 2008, mortgages become more likely to default. For group 2, the likelihood of default is consistently low.

Table 57 shows the results of the competing risks of prepayment and 90-days-delinquency based on model 3.30. The effect of each explanatory variable is very similar to the effect in the model using prepayment and default as the dependent variable, with the absolute value of the coefficients slightly changed. However, the dummy for the number of units has a positive effect on the 90-days-delinquency risk for the group 1 mortgages, the odds for mortgages covering more than one unit being delinquent instead of continued are about 1.058 times as high as the same odds for mortgages covering only one house unit.

[insert Table 57 here]

Figures 32 and 33 shows the comparison of the average conditional hazard for prepayment with and without controls for unobserved heterogeneity. The average conditional hazards for prepayment in figure 32 are based on the model using prepayment and default as dependent variables, and those in figure 33 are based on the model using prepayment and 90-days-delinquency as dependent variables.

Table 57. The monthly results for Phoenix based on Sueyoshi's proportional hazards model with prepayment and 90-days-delinquency as dependent variables (Baseline is loan age, P_1 =0.783, P_2 = 0.217)

	Group 1 Mortgages		Group 2 Mortgages	
	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.964	--	-0.313	--
	(0.292)		(0.581)	
Call option variables				
Call option part 1	5.621	--	29.364	--
	(0.811)		(5.607)	
Call option part 2	6.996	--	77.043	--
	(0.870)		(16.630)	
Call option part 3	4.968	--	35.353	--
	(0.782)		(2.739)	
Call option part 4	0.697	--	0.874	--
	(0.575)		(1.098)	
The unemployment rate variables				
The unemployment rate part 1	--	0.223	--	0.050
		(0.070)		(0.005)
The unemployment rate part 2	--	0.152	--	0.118
		(0.045)		(0.005)
Negative equity variables				
Negative equity part 1	--	0.020	--	-0.041
		(0.013)		(0.005)
Negative equity part 2	--	0.023	--	-0.020
		(0.008)		(0.005)
Negative equity part 3	--	0.019	--	-0.025
		(0.005)		(0.005)
Negative equity part 4	--	0.000	--	-0.028
		(0.003)		(0.005)
The negative equity dummy	--	0.870	--	0.775
		(0.190)		(0.005)
Original loan-to-value variables				
Original LTV part 1	--	2.124	--	3.842
		(0.820)		(0.005)
Original LTV part 2	--	0.678	--	1.444
		(0.838)		(0.005)
Original LTV part 3	--	-2.399	--	-7.508
		(3.445)		(0.005)
Original LTV part 4	--	2.974	--	2.804
		(0.675)		(0.005)
Original LTV part 5	--	-0.687	--	2.904
		(0.956)		(0.005)
Log loan size variables				
Log loan size part 1	0.554	0.173	1.879	-0.307
	(0.092)	(0.242)	(0.218)	(0.014)

Log loan size part 2	0.651	0.037	1.935	-0.562
	(0.139)	(0.326)	(0.414)	(0.005)
Log loan size part 3	-0.102	-0.842	1.470	-0.115
	(0.140)	(0.351)	(0.434)	(0.005)
Log loan size part 4	1.117	0.226	0.540	-0.037
	(0.111)	(0.241)	(0.303)	(0.005)
Loan age variables				
Loan age part 1	0.126	0.233	0.126	0.233
	(0.007)	(0.045)	(0.007)	(0.045)
Loan age part 2	0.010	0.036	0.010	0.036
	(0.004)	(0.013)	(0.004)	(0.013)
Loan age part 3	0.001	0.003	0.001	0.003
	(0.002)	(0.005)	(0.002)	(0.005)
Loan age part 4	0.005	0.002	0.005	0.002
	(0.001)	(0.003)	(0.001)	(0.003)
The credit score variables				
The credit score part 1	0.406	-10.047	6.350	-6.664
	(0.654)	(1.360)	(1.566)	(0.005)
The credit score part 2	1.693	-11.220	11.340	-7.110
	(0.931)	(1.947)	(1.514)	(0.005)
The credit score part 3	0.699	-6.259	2.236	-2.262
	(0.866)	(2.836)	(1.929)	(0.005)
The credit score part 4	3.920	-24.266	18.213	-24.822
	(0.942)	(5.163)	(2.888)	(0.005)
The debt-to-income ratio variables				
DTI part 1	-1.411	2.097	2.004	0.950
	(0.430)	(1.433)	(1.388)	(0.005)
DTI part 2	1.436	4.129	-0.037	3.332
	(0.563)	(1.343)	(1.354)	(0.005)
DTI part 3	-0.595	1.651	-0.779	0.619
	(0.612)	(1.250)	(1.458)	(0.005)
DTI part 4	0.480	1.200	-0.227	0.778
	(0.392)	(0.745)	(0.873)	(0.005)
The dummy for the number of units	-0.333	0.056	-1.532	-0.443
	(0.279)	(0.587)	(0.646)	(0.005)
Time period variables				
Time period part 1	0.020	0.001	0.013	-0.952
	(0.002)	(0.002)	(0.007)	(0.005)
Time period part 2	-0.024	0.014	-0.139	-0.098
	(0.001)	(0.008)	(0.007)	(0.005)
Time period part 3	0.000	0.012	0.139	-0.037
	(0.002)	(0.007)	(0.009)	(0.005)
Prepayment dummy	-12.341	--	-36.644	--
	(1.125)		(2.494)	
Default dummy	--	-8.046	--	-12.598
		(2.920)		(0.005)

Number of mortgages 12,734

[insert Figures 32 through 35 here]

Figures 32 and 33 present very similar results. When the loan age is below 79 months, the average conditional hazard for the prepayment is similar across the models with and without controls for unobserved heterogeneity. When the loan age is between 79 and 93 months, between 103 and 118 months and above 165 months, the model that does not control for unobserved heterogeneity underpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity. Moreover, when the loan age is between 120 and 165 months, the model that does not control for unobserved heterogeneity is overpredicts the hazard compared with the model that controls for unobserved heterogeneity.

Figure 34 shows the comparison of the average conditional hazard for default with and without unobserved heterogeneity. It shows that, when the loan age is below 49 months, the average conditional hazard for default is similar across the models with and without controls for unobserved heterogeneity. When the loan age is between 49 and 89 months, the model that does not control for unobserved heterogeneity overpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity. When the loan age is above 102 months, the model that does not control for unobserved heterogeneity slightly underpredicts the average conditional hazard.

Figure 35 compares the average conditional hazard for 90-days-delinquency with and without unobserved heterogeneity. It shows that, when the loan age is below 48 months, the average conditional hazard for 90-days-delinquency is similar across the models with and without unobserved heterogeneity. When the loan age is between 48 and 89 months, the model that does not control for unobserved heterogeneity overpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity. When the loan age is above

Figure 32. The average conditional hazard for prepayment for the competing risks model based on the proportional hazards model (prepayment and default as dependent variables)

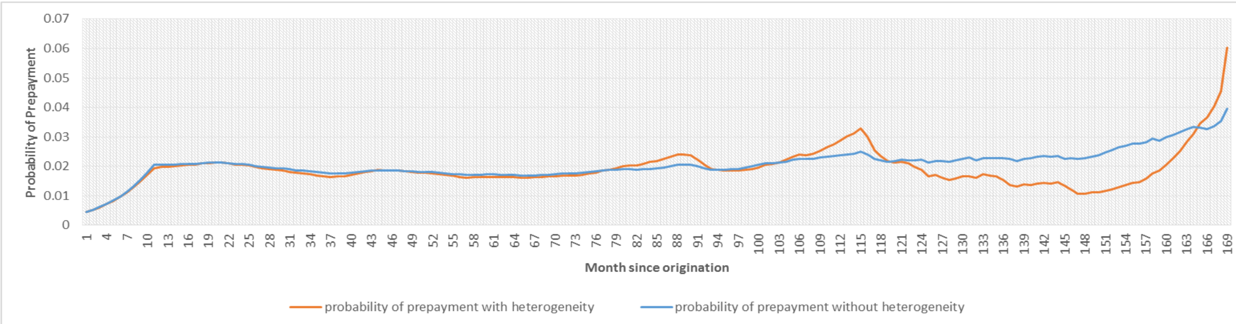


Figure 33. The average conditional hazard for prepayment for the competing risks model based on the proportional hazards model (prepayment and 90-days-delinquency as dependent variables)

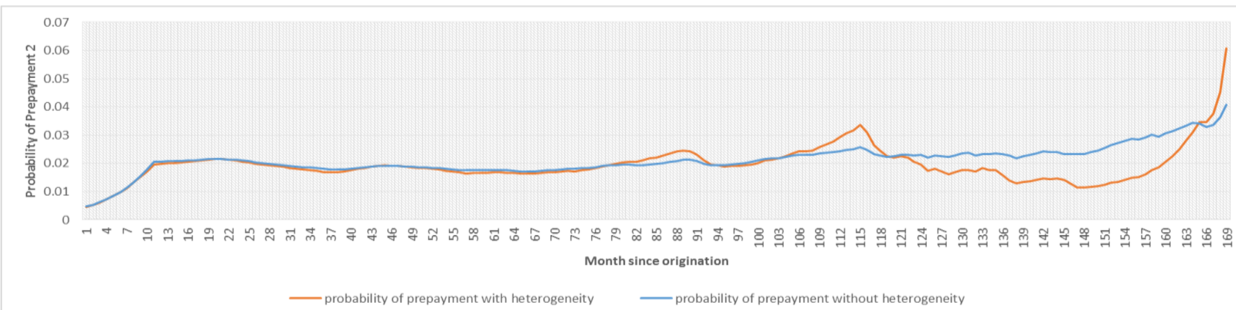


Figure 34. The average conditional hazard for default for the competing risks model based on the proportional hazards model (prepayment and default as dependent variables)

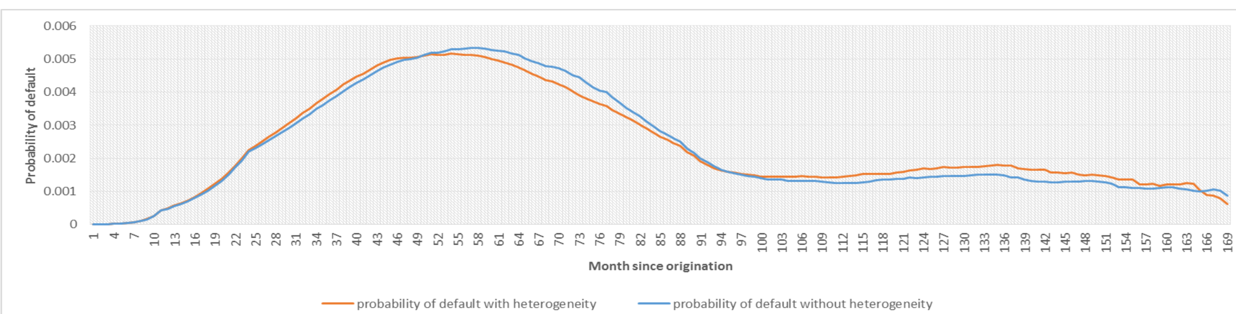
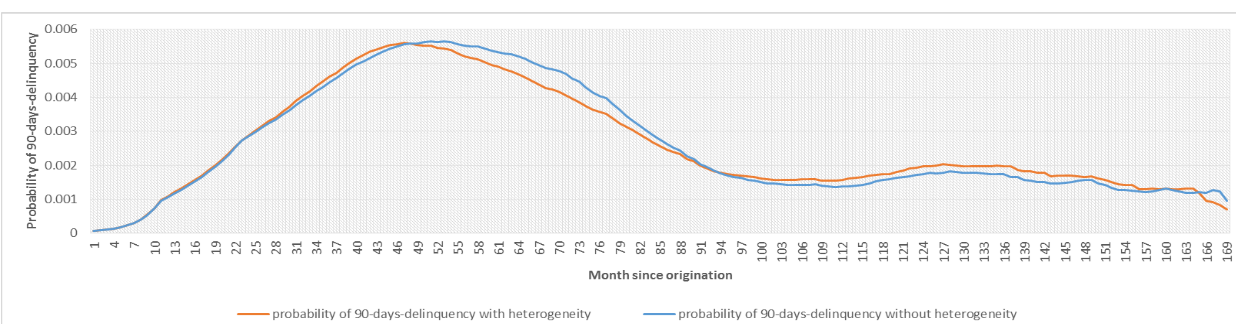


Figure 35. The average conditional hazard for 90-days-delinquency for the competing risks model based on the proportional hazards model (prepayment and 90-days-delinquency as dependent variables)



95 months, the model that does not control for unobserved heterogeneity slightly underpredicts the average conditional hazard.

Results 3: The results of the competing risks model with latent classes based on a class of transformation survival models

Table 58 shows the results of the competing risks of prepayment and default based on model 3.42. The results estimated by this model are very similar to the results given by model 3.30 because the coefficients of the prepayment risk for the group 1 mortgages and the coefficients of both risks for the group 2 mortgages are estimated by the proportional hazards framework. In the following part, this paper presents the results, which are different from those given by model 3.30.

[insert Table 58 here]

Based on the results in table 58, the probability of a mortgage belonging to the group 1 individuals is 0.772 and belonging to the group 2 individuals is 0.228, and the best transformation parameter C_d for the group 1 mortgages is 73.807.

For groups 1 and 2 individuals, the effect of the indicator for a prepayment penalty is negative. The odds for mortgages with a prepayment penalty are 0.754 times as high as the same odds for mortgages without a prepayment penalty for group 2 mortgages.

The effect of the value of the call option is all positive and the coefficients for the group 2 mortgages are larger than those for the group 1 mortgages. The value of the coefficients for group 2 mortgages is slightly higher than those given by model 3.30.

Table 58. The monthly results for Phoenix based on a Class of Transformation Survival Models for group 1 default and Sueyoshi's proportional hazards model for group 2 default and for group 1 and group 2 prepayment (Baseline is loan age, $P_1 = 0.772, P_2 = 0.228, C_d$ for group 1 is 73.807)

	Group 1 Mortgages		Group 2 Mortgages	
	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.979	--	-0.282	--
	(0.273)		(0.340)	
Call option variables				
Call option part 1	5.693	--	28.099	--
	(0.339)		(0.640)	
Call option part 2	6.938	--	71.590	--
	(0.320)		(0.606)	
Call option part 3	4.590	--	32.600	--
	(0.253)		(0.440)	
Call option part 4	0.231	--	0.756	--
	(0.367)		(0.619)	
The unemployment rate variables				
The unemployment rate part 1	--	0.102	--	0.051
		(0.110)		(0.005)
The unemployment rate part 2	--	0.045	--	0.118
		(0.081)		(0.005)
Negative equity variables				
Negative equity part 1	--	0.022	--	-0.041
		(0.018)		(0.005)
Negative equity part 2	--	0.035	--	-0.020
		(0.013)		(0.005)
Negative equity part 3	--	0.021	--	-0.025
		(0.009)		(0.005)
Negative equity part 4	--	0.007	--	-0.028
		(0.006)		(0.005)
The negative equity dummy	--	1.297	--	0.775
		(0.254)		(0.005)
Original loan-to-value variables				
Original LTV part 1	--	3.683	--	3.843
		(0.235)		(0.005)
Original LTV part 2	--	3.489	--	1.444
		(1.018)		(0.005)
Original LTV part 3	--	-15.940	--	-7.508
		(6.842)		(0.005)
Original LTV part 4	--	6.107	--	2.804
		(0.397)		(0.005)
Original LTV part 5	--	1.351	--	2.904
		(1.939)		(0.005)
Log loan size variables				
Log loan size part 1	0.525	0.612	1.805	-0.305
	(0.043)	(0.120)	(0.055)	(0.005)

Log loan size part 2	0.635	-1.108	1.913	-0.562
	(0.134)	(0.350)	(0.310)	(0.005)
Log loan size part 3	-0.164	0.014	1.538	-0.115
	(0.160)	(0.201)	(0.313)	(0.005)
Log loan size part 4	1.168	0.307	0.483	-0.037
	(0.117)	(0.333)	(0.277)	(0.005)
Loan age variables				
Loan age part 1	0.126	0.393	0.126	0.393
	(0.007)	(0.110)	(0.007)	(0.110)
Loan age part 2	0.010	0.091	0.010	0.091
	(0.003)	(0.019)	(0.003)	(0.019)
Loan age part 3	0.000	0.009	0.000	0.009
	(0.002)	(0.008)	(0.002)	(0.008)
Loan age part 4	0.005	0.000	0.005	0.000
	(0.001)	(0.004)	(0.001)	(0.004)
The credit score variables				
The credit score part 1	0.409	-15.731	8.491	-6.664
	(0.438)	(1.223)	(0.531)	(0.005)
The credit score part 2	1.656	-16.662	11.306	-7.110
	(0.397)	(1.401)	(0.458)	(0.005)
The credit score part 3	0.873	-3.024	1.659	-2.262
	(0.393)	(1.002)	(0.485)	(0.005)
The credit score part 4	4.287	-33.435	17.916	-24.822
	(0.519)	(4.098)	(0.133)	(0.005)
The debt-to-income ratio variables				
DTI part 1	-1.430	-0.554	1.951	0.950
	(0.350)	(1.161)	(0.478)	(0.005)
DTI part 2	1.495	6.671	-0.692	3.332
	(0.287)	(1.041)	(0.537)	(0.005)
DTI part 3	-0.646	0.274	-0.410	0.619
	(0.271)	(0.876)	(0.637)	(0.005)
DTI part 4	0.370	2.632	0.015	0.778
	(0.364)	(0.606)	(0.711)	(0.005)
The dummy for the number of units	-0.303	-1.008	-1.466	-0.443
	(0.245)	(0.522)	(0.106)	(0.005)
Time period variables				
Time period part 1	0.021	-0.007	0.014	-0.952
	(0.002)	(0.003)	(0.006)	(0.005)
Time period part 2	-0.024	0.083	-0.138	-0.098
	(0.001)	(0.015)	(0.006)	(0.005)
Time period part 3	-0.001	-0.005	0.140	-0.037
	(0.002)	(0.012)	(0.008)	(0.005)
Prepayment dummy	-12.011	--	-36.861	--
	(0.371)		(0.396)	
Default dummy	--	-11.871	--	-12.598
		(0.445)		(0.005)

Number of mortgages 12,734

The effect of the monthly unemployment rate for both groups is positive. For the group 1 mortgages, when the unemployment rate is below 5.700, the odds of default relative to continuity increase by 1.107 times with a 1 percent increase in the unemployment rate. When the rate is above 5.700, the odds of default relative to continuity increase by 1.046 times with a 1 percent increase in the unemployment rate. The results using model 3.42 for the group 1 mortgages are lower than those using model 3.30.

The effect of negative equity on the default risk is consistently positive for the group 1 mortgages; however, the effect is insignificantly negative for the group 2 mortgages. Therefore, the negative equity splines are not a strong factor for the group 2 mortgages. Moreover, the effect of negative equity dummy for both groups is significantly positive, which supports that when the put option is "in the money," households have more incentive to default their mortgages. For the group 1 mortgages, the odds for mortgages with negative equity being defaulted instead of continued are 3.658 times as high as the same odds for mortgages with non-negative equity. The result is higher than that from model 3.30

The results of original loan-to-value partially support the previous literature that the original loan-to-value has a positive effect on the default risk. When its value ranges between 0.780 and 0.800, the effect becomes insignificantly negative for both groups.

Consistent with the results in previous studies, the credit score has a strong negative effect on the default risk for both groups of mortgages. Moreover, the results indicate it has a strongly positive effect on the prepayment decision for both groups of mortgages. The results also indicate that both termination risks are sensitive to the high-level credit score which means that the households with high credit scores are more likely to prepay and less likely to default.

Moreover, compared with the results by model 3.30, the absolute value of the coefficients in this model is much higher.

The effect of the debt-to-income ratio on the default risk is significantly positive for group 2 mortgages, which support the hypothesis that household with high debt-to-income ratio are more likely to default. However, its effect on the default risk is insignificant for group 1 mortgages. Moreover, for the prepayment risk, the effect of the debt-to-income ratio is inconsistent for both groups.

The dummy for the number of units has a negative effect on both the prepayment and default risks for both groups of mortgages. For group 1, the odds for mortgages covering more than one house unit being prepaid instead of continued are about 0.738 times as high as the same odds for mortgages covering only one house unit. Moreover, the odds for mortgages covering more than one house unit being defaulted instead of continued are about 0.365 times as high as the same odds for mortgages covering only one house unit. Compared with the results given by model 3.30, the effect of the dummy for the number of units is slightly higher.

Log loan size has a significantly positive effect on the prepayment. However, the effect is negative when the loan size ranges between $149,941 and $210,029 for the group 1 mortgages. On the other hand, its effect on the default is positive for the group 1 mortgages and is negative for the group 2 mortgages.

As the baseline, the loan age has the same effect for both groups of mortgages. The effect of loan age for both risks in this model is positive and is slightly lower than that in model 3.30 for default risk. Moreover, the effect of both risks is the strongest when loan age is below 11 months.

The paper also controls for the time period; before January 2008, group 1 mortgages are less likely to default, from January 2008 to June 2010 mortgages are more likely to default. After June 2010, the likelihood of default decreases again.

Table 59 shows the results of the competing risks of prepayment and 90-days-delinquency based on model 3.42. The effect of each explanatory variable is very similar to the effect in the model using prepayment and default as the dependent variable, with the absolute value of the coefficients changes a little bit. However, the dummy for the number of units has a positive effect on the 90-days-delinquency risk for group 1 mortgages, the odds for mortgages covering more than one unit being delinquent instead of continued are about 3.865 times as high as the same odds for mortgages covering only one house unit.

[insert Table 59 here]

Figures 36 and 37 shows the comparison of the average conditional hazard for prepayment with and without controls for unobserved heterogeneity. The average conditional hazards for prepayment in figure 36 are based on the model using prepayment and default as dependent variables, and those in figure 37 are based on the model using prepayment and 90-days-delinquency as dependent variables.

[insert Figures 36 through 39 here]

Figures 36 and 37 present very similar results. When the loan age is below 79 months, the average conditional hazard for the prepayment is similar across the models with and without controls for unobserved heterogeneity. When the loan age is between 79 and 93 months, between 103 and 118 months and above 165 months, the model that does not control for unobserved heterogeneity underpredicts the average conditional hazard compared with the model that

Table 59. The monthly results for Phoenix based on a Class of Transformation Survival Models for group 1 90-days-deliqneucy and Sueyoshi's proportional hazards model for group 2 90-days-deliqneucy and for group 1 and group 2 prepayment (Baseline is loan age, P_1 =0.786, P_2 = 0.214, C_d for group 1 is 78.579)

	Group 1 Mortgages		Group 2 Mortgages	
	Prepayment	Default	Prepayment	Default
Prepayment penalty	-0.964	--	-0.323	--
	(0.292)		(0.632)	
Call option variables				
Call option part 1	5.629	--	29.659	--
	(0.835)		(4.418)	
Call option part 2	7.053	--	76.627	--
	(0.901)		(13.178)	
Call option part 3	5.018	--	35.634	--
	(0.818)		(2.976)	
Call option part 4	0.625	--	0.882	--
	(0.605)		(1.183)	
The unemployment rate variables				
The unemployment rate part 1	--	0.197	--	0.050
		(0.082)		(0.005)
The unemployment rate part 2	--	0.195	--	0.118
		(0.079)		(0.005)
Negative equity variables				
Negative equity part 1	--	0.027	--	-0.041
		(0.019)		(0.005)
Negative equity part 2	--	0.036	--	-0.020
		(0.013)		(0.005)
Negative equity part 3	--	0.039	--	-0.025
		(0.010)		(0.005)
Negative equity part 4	--	0.006	--	-0.028
		(0.006)		(0.005)
The negative equity dummy	--	1.109	--	0.775
		(0.268)		(0.005)
Original loan-to-value variables				
Original LTV part 1	--	2.416	--	3.842
		(0.932)		(0.005)
Original LTV part 2	--	1.973	--	1.444
		(0.987)		(0.005)
Original LTV part 3	--	-15.520	--	-7.508
		(0.813)		(0.005)
Original LTV part 4	--	5.035	--	2.804
		(1.020)		(0.005)
Original LTV part 5	--	7.549	--	2.904
		(9.626)		(0.005)
Log loan size variables				
Log loan size part 1	0.570	0.314	1.804	-0.309
	(0.092)	(0.338)	(0.274)	(0.012)

Log loan size part 2	0.629	-0.373	2.010	-0.562
	(0.145)	(0.539)	(0.435)	(0.005)
Log loan size part 3	-0.083	-0.559	1.409	-0.115
	(0.153)	(0.557)	(0.461)	(0.005)
Log loan size part 4	1.128	0.081	0.513	-0.037
	(0.114)	(0.401)	(0.307)	(0.005)
Loan age variables				
Loan age part 1	0.126	0.233	0.126	0.233
	(0.007)	(0.049)	(0.007)	(0.049)
Loan age part 2	0.010	0.049	0.010	0.049
	(0.004)	(0.017)	(0.004)	(0.017)
Loan age part 3	0.001	0.007	0.001	0.007
	(0.002)	(0.007)	(0.002)	(0.007)
Loan age part 4	0.004	0.001	0.004	0.001
	(0.001)	(0.004)	(0.001)	(0.004)
The credit score variables				
The credit score part 1	0.375	-16.702	6.599	-6.664
	(0.666)	(2.162)	(1.607)	(0.005)
The credit score part 2	1.589	-23.037	11.816	-7.110
	(0.934)	(2.657)	(1.341)	(0.005)
The credit score part 3	0.859	-5.375	1.675	-2.262
	(0.801)	(2.724)	(1.564)	(0.005)
The credit score part 4	3.950	-34.688	18.317	-24.822
	(0.874)	(1.723)	(3.324)	(0.005)
The debt-to-income ratio variables				
DTI part 1	-1.465	1.309	2.262	0.950
	(0.434)	(1.988)	(1.417)	(0.005)
DTI part 2	1.448	6.870	-0.182	3.332
	(0.549)	(1.801)	(1.471)	(0.005)
DTI part 3	-0.648	2.076	-0.533	0.619
	(0.566)	(1.744)	(1.445)	(0.005)
DTI part 4	0.474	1.976	-0.265	0.778
	(0.392)	(1.191)	(0.958)	(0.005)
The dummy for the number of units	-0.251	1.352	-1.614	-0.443
	(0.267)	(0.942)	(0.645)	(0.005)
Time period variables				
Time period part 1	0.020	0.001	0.014	-0.953
	(0.002)	(0.003)	(0.007)	(0.005)
Time period part 2	-0.025	0.033	-0.140	-0.098
	(0.001)	(0.014)	(0.007)	(0.005)
Time period part 3	0.000	0.003	0.141	-0.037
	(0.002)	(0.012)	(0.009)	(0.005)
Prepayment dummy	-12.493	--	-36.019	--
	(1.133)		(2.823)	
Default dummy	--	-5.325	--	-12.598
		(4.007)		(0.005)

Number of mortgages 12,734

Figure 36. The average conditional hazard for prepayment for the competing risks model based on a class of transformation survival models (prepayment and default as dependent variables)

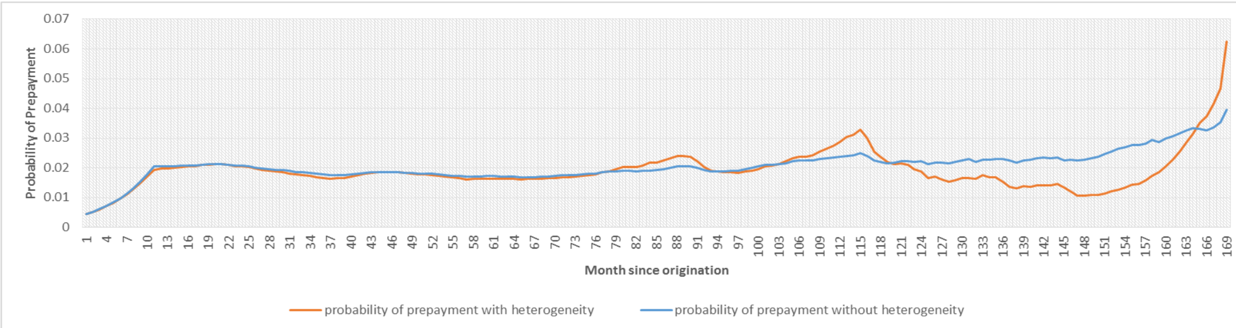


Figure 37. The average conditional hazard for prepayment for the competing risks model based on a class of transformation survival models (prepayment and 90-days-delinquency as dependent variables)

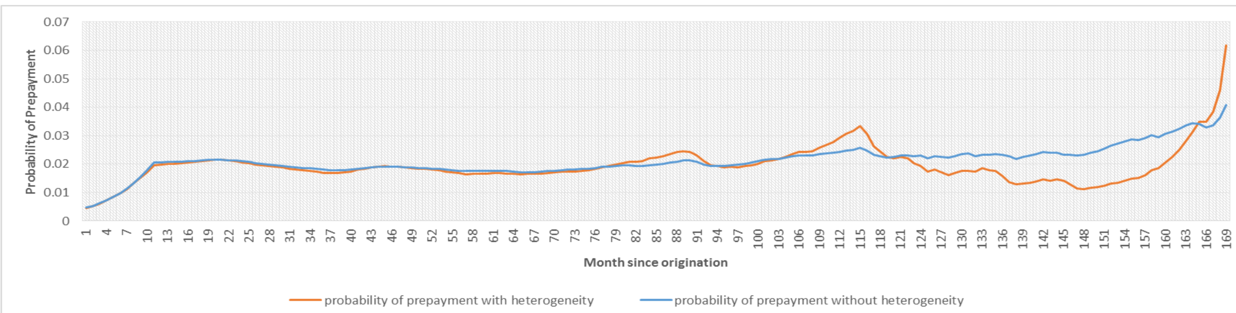


Figure 38. The average conditional hazard for default for the competing risks model based on a class of transformation survival models (prepayment and default as dependent variables)

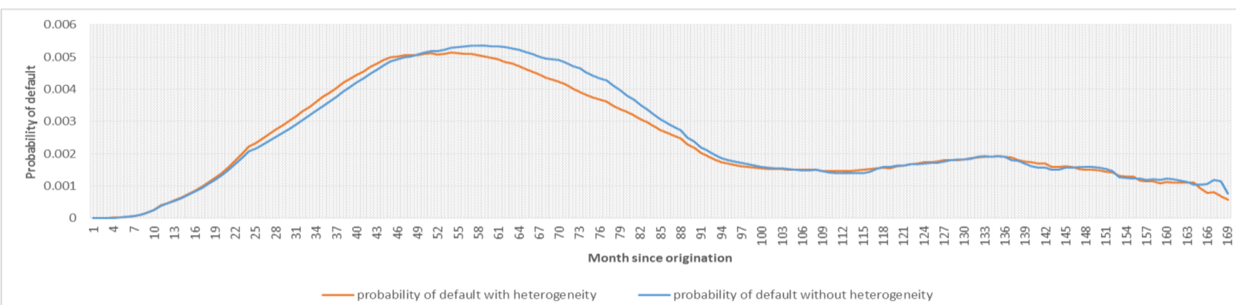
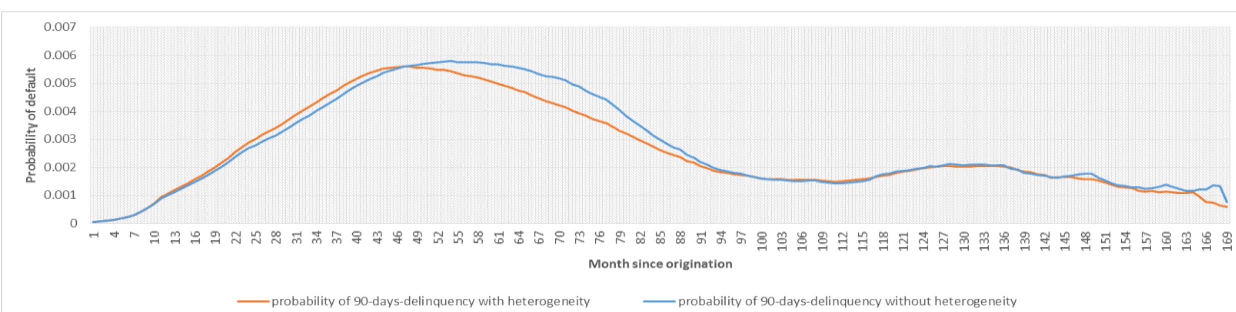


Figure 39. The average conditional hazard for 90-days-delinquency for the competing risks model based on a class of transformation survival models (prepayment and 90-days-delinquency as dependent variables)



controls for unobserved heterogeneity. Moreover, when the loan age is between 120 and 165 months, the model that does not control for unobserved heterogeneity is highly overpredicts the hazard compared with the model that controls for unobserved heterogeneity.

Figure 38 shows the comparison of the average conditional hazard for default with and without unobserved heterogeneity. It shows that, when the loan age is between 50 and 96 months, the model that does not control for unobserved heterogeneity overpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity. The average conditional hazard for default is similar across two models in all other ranges.

Figure 39 compares the average conditional hazard for 90-days-delinquency with and without unobserved heterogeneity. It shows that, when the loan age is between 49 and 92 months, the model that does not control for unobserved heterogeneity overpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity. The average conditional hazard for 90-days-delinquency is similar across two models in all other ranges.

Simulation of the Stable Housing Price

The Case-Shiller index in Figure 40 shows that, starting from around September 2004, the housing price index dramatically increases from 141.64 to 227.01 in August 2007, and then sharply decreases to 103.56 in May 2010. The boom and bust of the housing price becomes the main trigger of the financial crisis that happened in around 2008 and the default rate of the mortgages remains high during the crisis. This paper simulates what the average conditional default hazard and the average conditional 90-days-delinquency hazard when the housing price remains the same since September 2004.

[insert Figure 40 here]

Assuming that the housing price stops changing after September 2004, let CS_N indicate the new Case-Shiller index and CS_0 indicate the original Case-Shiller index. Before September 2004, the new index is equal to the original Case-Shiller index ($CS_N = CS_0$) and after September 2004, the index remains the same as the one in September 2004 ($CS_N = CS_{0\ at\ sep\ 2004} = 141.64$). Therefore, the new value of the property will change to $V_N = V_0 \frac{CS_N}{CS_0}$. To keep the original loan-to-value unchanged, the new loan size $L_N = L_0 \frac{CS_N}{CS_0}$. When the loan size changes, the debt-to-income ratio also changes to $DTI_{N\ new} = DTI_0 \frac{CS_N}{CS_0}$. Because the negative equity is calculated as the difference between the remaining balance and the property value in each month, the change of the loan size and property value leads to the change of the negative equity. The detailed calculation of the new negative equity can be found in Appendix D. After estimating the average conditional default hazard and the average conditional 90-days-delinquency hazard based on the new dataset by the above three competing risks models, the results are compared with those calculated by the original dataset.

[insert Figures 41 through 43 here]

Figures 41, 42 and 43 show the comparison results of the average conditional hazard for default estimated by three models. The results indicate that, in the case with the actual housing prices, the estimated average conditional default hazard increases dramatically from 0.04 percent when the age month is 11 to the maximum 0.51 percent when the age month is 54, and then sharply decreases to 0.17 percent when the age month is 93. After that, the average conditional default hazard slowly changes until the age month reaches 169. In the case where the housing price is constant after September 2004, the simulated average conditional hazard increases

Figure 40. Monthly Case-Shiller Index

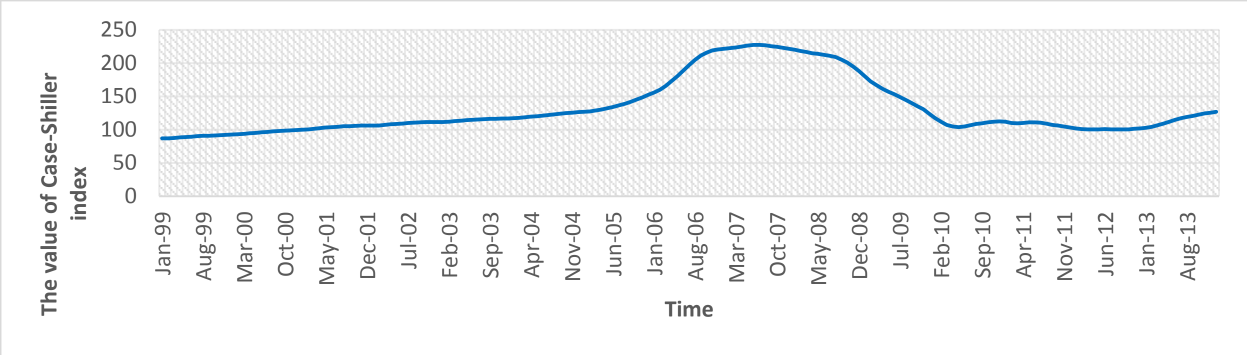


Figure 41. The comparison of simulated and non-simulated average conditional default hazards for the competing risks model based on the multinomial logit

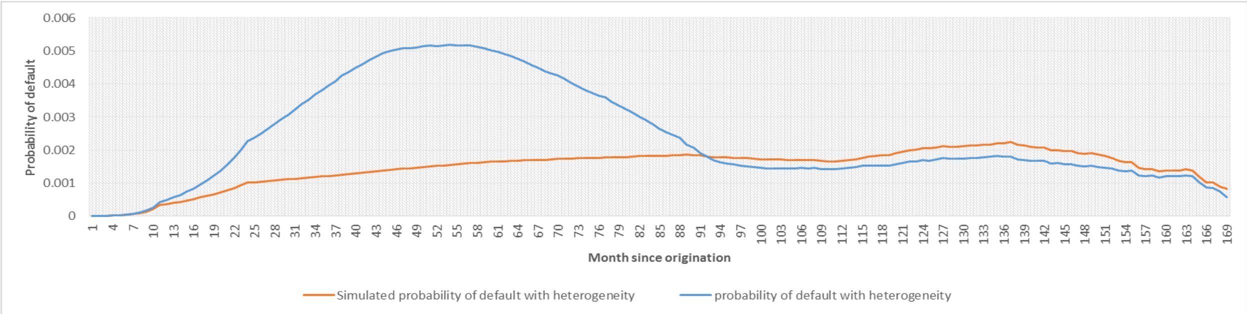


Figure 42. The comparison of simulated and non-simulated average conditional default hazards for the competing risks model based on the proportional hazards model

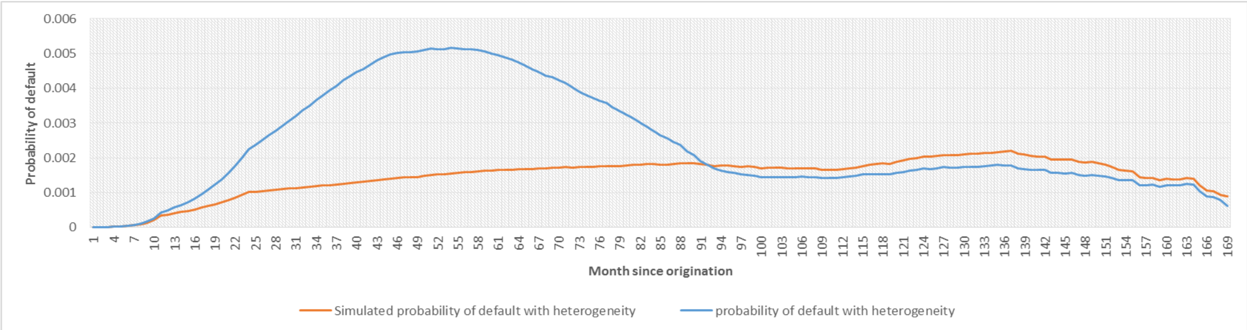
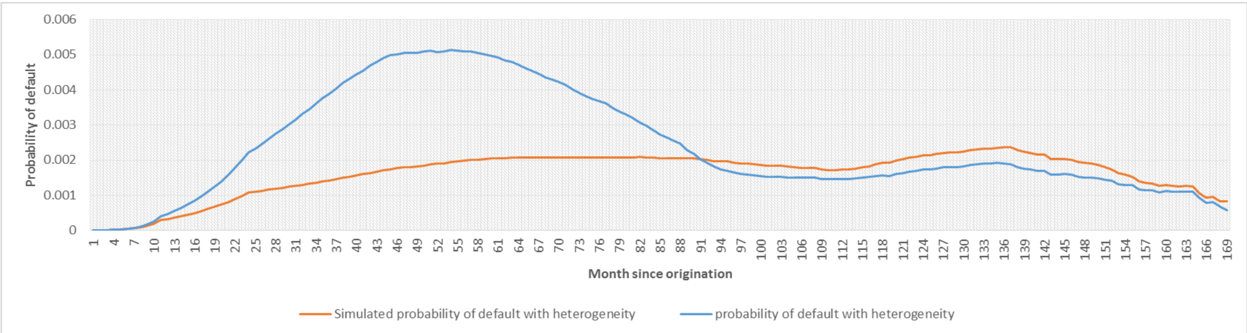


Figure 43. The comparison of simulated and non-simulated average conditional default hazards for the competing risks model based on a class of transformation survival models



slowly from age month 1 up to age month 169 with an average increase rate of 3.72 percent. The figures clearly show that the sharp increase and decrease of the default hazard between age month 10 and 93 is deleted by keeping the housing price unchanged and the average difference of the conditional hazard is 0.21 percent across two cases.

The comparison results of the conditional 90-days-delinquency hazard estimated by three models are presented in figures 44, 45 and 46. The results indicate that, in the case with the actual housing prices, the estimated average conditional 90-days-delinquency hazard increases dramatically from 0.05 percent when the age month is 9 to the maximum 0.56 percent when the age month is 47, and then sharply decreases to 0.18 percent when the age month is 94. After that, the average conditional 90-days-delinquency hazard slowly changes until the age month reaches 169. In the case where the housing price is constant after September 2004, the simulated average conditional hazard increases slowly from age month 1 up to age month 169 with an average increase rate of 1.80 percent. The figures clearly show the sharp increase and decrease of the 90-days-delinquency hazard between age month 9 and 94 is deleted by keeping the housing price unchanged and the average difference of the conditional hazard is 0.20 percent across two cases.

[insert Figures 44 through 46 here]

Summary

The results of the three models controlling unobserved heterogeneity support most of the hypothesis made by the previous literature. For the three models, the probability that a mortgage belongs to group 1 individuals is around 0.766 and the probability that a mortgage belongs to group 2 individuals is around 0.234.

Figure 44. The comparison of simulated and non-simulated average conditional 90-days-delinquency hazards for the competing risks model based on the multinomial logit

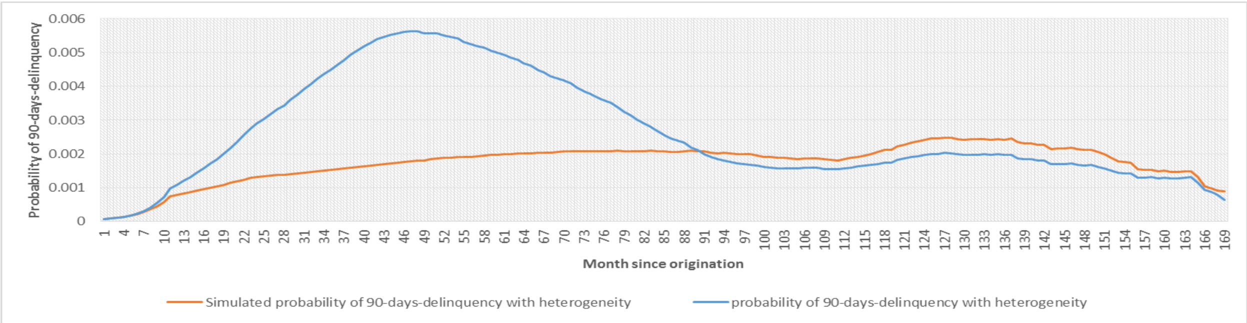


Figure 45. The comparison of simulated and non-simulated average conditional 90-days-delinquency hazards for the competing risks model based on the proportional hazards model

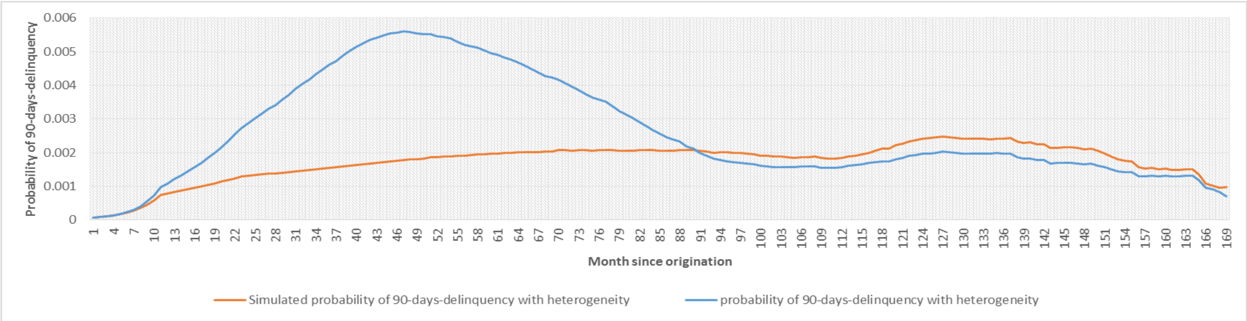
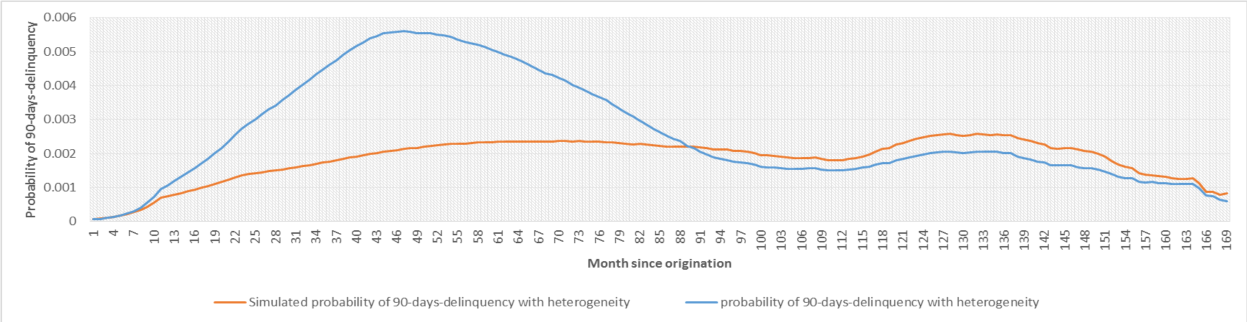


Figure 46. The comparison of simulated and non-simulated average conditional 90-days-delinquency hazards for the competing risks model based on a class of transformation survival models



Compared the coefficients across groups, the major differences are summarized in the following. The indicator for a prepayment penalty has a negative on the prepayment risk for both groups and its effect on the group 1 mortgages is stronger. The value of call option has a positive effect on the prepayment risk and it has a lager effect on the group 2 mortgages. The effect of unemployment rate is positive and it has a larger effect on group 1 mortgages when the rate is below 5.7 and a larger effect on group 2 mortgages when the rate is above 5.7. The negative equity dummy has a positive effect on the default/90-days-delinquency risk and its effect on the group 1 mortgages is larger. The original loan-to-value has a stronger positive effect on the 90-days-delinquency for the group 2 mortgages. The credit score has a positive effect on the prepayment risk and its effect on the group 2 mortgages is stronger. The dummy for the number of units has a negative effect on both prepayment and default risks for both groups of mortgages. And the effect is stronger on prepayment risk for the group 2 mortgages. Moreover, it has a positive effect on the 90-days-delinquency risk for the group 1 mortgages.

The paper also compares the average conditional hazard for prepayment, default and 90-days-delinquency estimated by models with and without unobserved heterogeneity. The results show that when the loan age is younger than 79 months, the average conditional hazard for the prepayment is similar across the models with and without controls for unobserved heterogeneity. When the loan age is between 79 and 93 months, between 103 and 118 months and above 165 months, the model that does not control for unobserved heterogeneity underpredicts the average conditional prepayment hazard compared with the model that controls for unobserved heterogeneity. Moreover, when the loan age is between 120 and 165 months, the model that does not control for unobserved heterogeneity highly overpredicts the prepayment hazard compared with the model that controls for unobserved heterogeneity. For the average conditional default and 90-days-delinquency hazard, overall, when the loan age is between around 49 and 94 months, the model that does not control for unobserved heterogeneity overpredicts the average conditional hazard compared with the model that controls for unobserved heterogeneity. The average conditional hazard for default and 90-days-delinquency is very similar across the two models in all other ranges.

By assuming that housing prices remain constant after September 2004 and the original loan-to-value remains unchanged, the value of the property, the mortgage size, the debt-to-income ratio, the value of the negative equity and the negative equity dummy change accordingly. The simulated average conditional default and 90-days-delinquency hazards are compared with those estimated based on the real trend of the housing price. The results shows that, in the case when housing prices change across time, the average conditional hazard dramatically increases from around age month 11 and reaches the maximum hazard at around age month 54, and then sharply decreases until around age month 93. This dramatic change of the average conditional hazard trend disappears in the case when housing prices are assumed to be unchanged after September 2004. The simulated average conditional hazard slowly increases from age month 1 up to age month 169 with an average increase rate of 3.72 percent for default hazard and 1.80 percent for 90-days-delinquency hazard. The average difference of the conditional hazard is approximately 0.21 percent between the two cases.

Appendix A

The calendar-time prepayment rate is calculated as:

$$Pr_i = \frac{p_i}{sm_i}$$

in which Pr_i is the prepayment rate in month i, p_i is the number of prepayments in month i, and sm_i is the number of surviving mortgages in month i, which equals surviving mortgages in month $i-1$ plus the number of mortgages originated in month i minus the number of mortgages terminated in month i. When $i = 1$, sm_i equals the number of mortgages originated in month i minus the number of mortgages terminated in month i.

The calendar-time default rate is calculated as:

$$Dr_i = \frac{d_i}{sm_i}$$

in which Dr_i is the default rate in month i, d_i is the number of defaults in month i, and sm_i is the number of surviving mortgages in month i.

The value of the call option is computed as the ratio of the difference between the market value of the mortgage and the book value of the mortgage to the market value of the mortgage.

$$CallOption_t = \frac{Market\ Value_{tj} - Book\ Value_t}{Market\ Value_{tj}} = 1 - \frac{Book\ Value_t}{Market\ Value_{tj}}$$

in which $Book\ Value_t$ is the book value of a mortgage at age t months, which is calculated as:

$$Book\ Value_t = \sum_{i=1}^{360-t} \frac{Monthly\ Payment}{(1 + contract\ rate/12)^i}$$

and $Market\ Value_{tj}$ is the market value of a mortgage at age t months in calendar month j, which is calculated as:

$$Market\ Value_{tj} = \sum_{i=1}^{360-t} \frac{Monthly\ Payment}{(1 + market\ rate_j/12)^i}$$

When calculating book value, the contract rate is used, which is original interest rate in the dataset. The market interest rate used in a given calendar year j is calculated by the average note rate of mortgages originated in that year in the data.

Negative equity is an important determinant of default risk. Equity is negative when the property is worth less than the remaining balance. When equity is negative, the holder has more incentive to default to keep their wealth.

The functions used to calculate this are:

$$X_i = current\ property\ value_i - remaining\ balance_i$$

$$negative\ equity_i = \begin{cases} aboslute\ value\ of\ X_i & if X_i < 0 \\ 0 & if\ X_i \geq 0 \end{cases}$$

in which current property value at month i is calculated as:

$$current\ property\ value_i = \exp(\ln\left(\frac{Case\ Schiller\ index_i}{Case\ Schiller\ index_0}\right)) \times origianal\ property\ value$$

where $Case\ Schiller\ index_0$ is the Case-Shiller index at the note date of a mortgage and $Case\ Schiller\ index_i$ is the Case-Shiller index at month i. And the remaining balance at month i is calculated as:

$$remaining\ balance_i = \frac{original\ balance[\left(1+\frac{contract\ rate}{12}\right)^{360} - (1+\frac{contract\ rate}{12})^{i}]}{\left(1+\frac{contract\ rate}{12}\right)^{360} - 1}$$

Appendix B

The following process is constructing the competing risks model based on the duration model introduced by Prentice and Gloeckler (1978), Bruce D. Meyer (1990) and the specification discussed by Glenn T. Sueyoshi (1992). This competing risks model can be implemented in any discrete time duration analysis involving three choices in each time period; however, this study only focuses on analyzing the competing risks of prepayment and default in the single-family mortgage market.

According to the duration model introduced by Prentice and Gloeckler (1978) and Meyer (1990), two choices are involved in each time period—continue the event and terminate the event. Therefore, the likelihood of event i to be continued at time t is

$$L_t^C(\beta) = \prod_{s=1}^{t} \exp(-\exp(z_i(s)\beta))$$

$$= \exp(-\exp(z_i(1)\beta))\exp(-\exp(z_i(2)\beta)) \dots \exp(-exp(z_i(t)\beta))$$

$$= \exp(-\exp(z_i(1)\beta) - \exp(z_i(2)\beta) \dots - exp(z_i(t)\beta))$$

$$= \exp\left(-\sum_{s=1}^{t} \exp(z_i(s)\beta)\right)$$

$$= \exp\left(-\sum_{s=1}^{t} \exp(z_i(s)\beta)\right) - 0$$

$$= -\left[\exp(-\infty) - \exp\left(-\sum_{s=1}^{t}\exp(z_i(s)\beta)\right)\right]$$

$$= \int_{\sum_{s=1}^{t}\exp(z_i(s)\beta)}^{\infty} (-\exp(-u))'\,du$$

$$= \int_{\sum_{s=1}^{t}\exp(z_i(s)\beta)}^{\infty} \exp(-u)\,du \qquad \text{A. 1}$$

The likelihood of event i to be terminating at time t is

$$L_t^T(\beta) = (1 - \exp(-\exp(z_i(t)\beta)))\prod_{s=1}^{t-1}\exp(-\exp(z_i(s)\beta))$$

$$= \prod_{s=1}^{t-1}\exp(-\exp(z_i(s)\beta)) - \prod_{s=1}^{t}\exp(-\exp(z_i(s)\beta))$$

$$= \exp\left(-\sum_{s=1}^{t-1}\exp(z_i(s)\beta)\right) - \exp\left(-\sum_{s=1}^{t}\exp(z_i(s)\beta)\right)$$

$$= -\left[\exp\left(-\sum_{s=1}^{t}\exp(z_i(s)\beta)\right) - \exp\left(-\sum_{s=1}^{t-1}\exp(z_i(s)\beta)\right)\right]$$

$$= \int_{\sum_{s=1}^{t-1}\exp(z_i(s)\beta)}^{\sum_{s=1}^{t}\exp(z_i(s)\beta)} (-\exp(-u))'\,du$$

$$= \int_{\sum_{s=1}^{t-1}\exp(z_i(s)\beta)}^{\sum_{s=1}^{t}\exp(z_i(s)\beta)} \exp(-u)\,du$$

$$= \int_{\alpha_{t-1}}^{\alpha_t} f(u)du \qquad \text{A. 2}$$

in which $f(u) = \exp(-u)$, $\alpha_t = \sum_{s=1}^{t}\exp(z_i(s)\beta)$, and $\alpha_{t-1} = \sum_{s=1}^{t-1}\exp(z_i(s)\beta)$.

When there are three choices involved, continue the mortgage, prepayment and default, based on the specification in Sueyoshi's paper, the likelihood of continuing at time t is

$$l_t^C(\beta) = \int_{\alpha_t^1}^{\infty} \int_{\alpha_t^2}^{\infty} f(u_1, u_2) du_1 du_2$$

$$= \int_{\alpha_t^1}^{\infty} f(u_1) du_1 \int_{\alpha_t^2}^{\infty} f(u_2) du_2$$

$$= \int_{\sum_{s=1}^{t} \exp(z_{pi}(s)\beta)}^{\infty} \exp(-u_1)\, du_1 \int_{\sum_{s=1}^{t} \exp(z_{di}(s)\beta)}^{\infty} \exp(-u_2)\, du_2$$

$$= \exp\left(-\sum_{s=1}^{t} \exp(z_{pi}(s)\beta_p)\right) \exp\left(-\sum_{s=1}^{t} \exp(z_{di}(s)\beta_d)\right)$$

$$= \prod_{i=1}^{t} \exp\big(-\exp(z_{pi}(s)\beta_p)\big) \exp(-\exp(z_{di}(s)\beta_d)) \qquad \text{A. 3}$$

The likelihood of prepaying is

$$l_t^P(\beta) = \int_{\alpha_{t-1}^1}^{\alpha_t^1} \int_{\alpha_t^2}^{\infty} f(u_1, u_2) du_1 du_2$$

$$= \int_{\alpha_{t-1}^1}^{\alpha_t^1} f(u_1) du_1 \int_{\alpha_t^2}^{\infty} f(u_2) du_2$$

$$= \int_{\sum_{s=1}^{t-1} \exp(z_{pi}(s)\beta_p)}^{\sum_{s=1}^{t} \exp(z_{pi}(s)\beta_p)} \exp(-u_1)\, du_1 \int_{\sum_{s=1}^{t} \exp(z_{di}(s)\beta_d)}^{\infty} \exp(-u_2)\, du_2$$

$$= \left[\exp\left(-\sum_{s=1}^{t-1} \exp(z_{pi}(s)\beta_p)\right) - \exp\left(-\sum_{s=1}^{t} \exp(z_{pi}(s)\beta_p)\right)\right] \exp\left(-\sum_{s=1}^{t} \exp(z_{di}(s)\beta_d)\right)$$

$$= \left[1 - \exp\left(-\exp(z_{pi}(t)\beta_p)\right)\right] \exp\left(-\sum_{s=1}^{t-1} \exp(z_{pi}(s)\beta_p)\right) \exp\left(-\sum_{s=1}^{t} \exp(z_{di}(s)\beta_d)\right)$$

$$= [1 - \exp(-\exp(z_{pi}(t)\beta_p))] \exp(-\exp(z_{di}(t)\beta_d)) \prod_{s=1}^{t-1} \exp(-\exp(z_{pi}(s)\beta_p)) \exp(-\exp(z_{di}(s)\beta_d)) \quad \text{A. 4}$$

Similarly, the likelihood of defaulting is

$$l_t^D(\beta) = \int_{\alpha_{t-1}^2}^{\alpha_t^2} \int_{\alpha_t^1}^{\infty} f(u_1, u_2) du_1 du_2$$

$$= [1 - \exp(-\exp(z_{di}(t)\beta_d))] \exp(-\exp(z_{pi}(t)\beta_p)) \prod_{s=1}^{t-1} \exp(-\exp(z_{pi}(s)\beta_p)) \exp(-\exp(z_{di}(s)\beta_d)) \quad \text{A. 5}$$

Therefore, the likelihood function for a sample of N individuals is now given by:

$$l(\beta) = \prod_{S=1}^{N} l_t^P(\beta)^{\delta_p} l_t^D(\beta)^{\delta_d} l_t^C(\beta)^{\delta_c}$$

$$= \prod_{i=1}^{N} \{[1 - \exp(-\exp(z_{pi}(t)\beta_p))] \exp(-\exp(z_{di}(t)\beta_d))\}^{\delta_p}$$

$$\{[1 - \exp(-\exp(z_{di}(t)\beta_d))]\exp(-\exp(z_{pi}(t)\beta_p))\}^{\delta_d}$$

$$\{\exp(-\exp(z_{pi}(t)\beta_p)) \exp(-\exp(z_{di}(t)\beta_d))\}^{\delta_c} \prod_{s=1}^{t-1} \exp(-\exp(z_{pi}(s)\beta_p)) \exp(-\exp(z_{di}(s)\beta_d))$$

$$= \prod_{i=1}^{N} \prod_{s=1}^{t} \{[1 - \exp(-\exp(z_{pi}(s)\beta_p))] \exp(-\exp(z_{di}(s)\beta_d))\}^{\delta_p}$$

$$\{[1 - \exp(-\exp(z_{di}(s)\beta_d))] \exp(-\exp(z_{pi}(s)\beta_p))\}^{\delta_d} \{\exp(-\exp(z_{pi}(s)\beta_p)) \exp(-\exp(z_{di}(s)\beta_d))\}^{\delta_c} \quad A. 6$$

in which δ_p is the indicator for prepaying, $\delta_d$is the indicator for defaulting, and $\delta_c$is the indicator for continuing.

An important assumption for model 2.2 is that the time in the analysis is a discrete time rather than a grouped time. When a default/prepayment occurs in the time period t, the mortgage is treated as default/prepayment and is removed from the sample. One of the significant weakness of this assumption is that it does not consider the time period between the default/prepayment point and the time point $t+1$. Therefore, model 2.2 could lead to estimation bias.

One of the weaknesses of likelihood function A.6 is that it does not consider the time period between a prepayment or default and the final point in that period. Therefore, model A.6 could lead to estimation bias. Based on the specification in Sueyoshi (1992), a $g(u)$ function is involved to estimate the missing part. The new likelihood of prepaying is

$$L_t^P(\beta) = \int_{\alpha_{t-1}^1}^{\alpha_t^1} \int_{g(u_1)}^{\infty} f(u_1, u_2) du_1 du_2$$

$$= \int_{\alpha_{t-1}^1}^{\alpha_t^1} f(u_1) \int_{g(u_1)}^{\infty} f(u_2) du_2 \, du_1$$

$$= \int_{\sum_{s=1}^{t-1} \exp(z_{pi}(s)\beta_p)}^{\sum_{s=1}^{t} \exp(z_{pi}(s)\beta_p)} \exp(-u_1) \int_{g(u_1)}^{\infty} \exp(-u_2) \, du_2 du_1$$

$$= \int_{\sum_{s=1}^{t-1} \exp(z_{pi}(s)\beta_p)}^{\sum_{s=1}^{t} \exp(z_{pi}(s)\beta_p)} \exp(-u_1) \exp(-u_1 \frac{\exp(z_{di}(t)\beta_d)}{\exp\left(z_{pi}(t)\beta_p\right)} - \sum_{s=1}^{t-1} \exp(z_{di}(s)\beta_d)$$

$$+ \frac{\exp(z_{di}(t)\beta_d)}{\exp\left(z_{pi}(t)\beta_p\right)} \sum_{s=1}^{t-1} \exp\left(z_{pi}(s)\beta_p\right)) du_1$$

$$= \int_{\sum_{s=1}^{t-1}\exp(z_{pi}(s)\beta_p)}^{\sum_{s=1}^{t}\exp(z_{pi}(s)\beta_p)} \exp\left(-u_1(1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}) - \sum_{s=1}^{t-1}\exp\left(z_{di}(s)\beta_d\right)\right.$$

$$\left. + \frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right)\right) du_1$$

$$= \frac{1}{1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}}\left[\exp\left(-(1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)})\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right)\right.\right.$$

$$\left. - \sum_{s=1}^{t-1}\exp\left(z_{di}(s)\beta_d\right) + \frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right)\right)$$

$$- \exp\left(-(1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)})\sum_{s=1}^{t}\exp\left(z_{pi}(s)\beta_p\right) - \sum_{s=1}^{t-1}\exp\left(z_{di}(s)\beta_d\right)\right.$$

$$\left.\left. + \frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right)\right)\right]$$

$$= \frac{1}{1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}}\left(1-\exp\left(-\left(1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}\right)\exp\left(z_{pi}(t)\beta_p\right)\right)\right) \exp(-(1$$

$$+ \frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)})\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right))\exp(-\sum_{i=1}^{t-1}\exp\left(z_{di}(t)\beta_d\right)$$

$$+ \frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right))$$

$$= \frac{1}{1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}}\left(1-\exp\left(-\left(1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}\right)\exp\left(z_{pi}(t)\beta_p\right)\right)\right)\ \exp(-(1$$

$$+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)})\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right)-\sum_{s=1}^{t-1}\exp\left(z_{di}(s)\beta_d\right)$$

$$+\frac{\exp(z_{di}(t)\beta_d)}{\exp\left(z_{pi}(t)\beta_p\right)}\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right))$$

$$= \frac{1}{1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}}\left(1-\exp\left(-\left(1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}\right)\exp\left(z_{pi}(t)\beta_p\right)\right)\right)$$

$$\exp(-\sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right))\exp(-\sum_{s=1}^{t-1}\exp\left(z_{di}(s)\beta_d\right))$$

$$= \frac{1}{1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}}\left(1-\exp\left(-\left(1+\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}\right)\exp\left(z_{pi}(t)\beta_p\right)\right)\right)$$

$$\prod_{s=1}^{t-1}\exp\left(-\exp\left(z_{pi}(s)\beta_p\right)\right)\exp\left(-\exp\left(z_{di}(s)\beta_d\right)\right) \qquad A.7$$

in which the $g(u_1)$ function is:

$$g(u_1) = \alpha_{t-1}^2 + (u_1 - \alpha_{t-1}^1)\frac{\alpha_t^2-\alpha_{t-1}^2}{\alpha_t^1-\alpha_{t-1}^1}$$

$$= u_1\frac{\alpha_t^2-\alpha_{t-1}^2}{\alpha_t^1-\alpha_{t-1}^1} + \alpha_{t-1}^2 - \alpha_{t-1}^1\frac{\alpha_t^2-\alpha_{t-1}^2}{\alpha_t^1-\alpha_{t-1}^1}$$

$$= u_1\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)} + \sum_{s=1}^{t-1}\exp\left(z_{di}(s)\beta_d\right) - \sum_{s=1}^{t-1}\exp\left(z_{pi}(s)\beta_p\right)\frac{\exp\left(z_{di}(t)\beta_d\right)}{\exp\left(z_{pi}(t)\beta_p\right)}$$

Similarly, the new likelihood of defaulting is

$$L_t^D(\beta) = \int_{\alpha_{t-1}^2}^{\alpha_t^2} \int_{g(u_2)}^{\infty} f(u_1, u_2) du_1 du_2$$

$$= \frac{1}{1 + \frac{\exp(z_{pi}(t)\beta_p)}{\exp(z_{di}(t)\beta_d)}} \left(1 - \exp\left(-\left(1 + \frac{\exp(z_{pi}(t)\beta_p)}{\exp(z_{di}(t)\beta_d)}\right) \exp(z_{di}(t)\beta_d)\right)\right)$$

$$\prod_{s=1}^{t-1} \exp(-\exp(z_{pi}(s)\beta_p)) \exp(-\exp(z_{di}(s)\beta_d)) \qquad A.8$$

in which the $g(u_2)$ function is:

$$g(u_2) = \alpha_{t-1}^1 + (u_2 - \alpha_{t-1}^2) \frac{\alpha_t^1 - \alpha_{t-1}^1}{\alpha_t^2 - \alpha_{t-1}^2}$$

$$= u_S \frac{\exp(z_{pi}(t)\beta_p)}{\exp(z_{di}(t)\beta_d)} + \sum_{i=1}^{t-1} \exp(z_{pi}(s)\beta_p) - \sum_{s=1}^{t-1} \exp(z_{di}(s)\beta_d) \frac{\exp(z_{pi}(t)\beta_p)}{\exp(z_{di}(t)\beta_d)}$$

The likelihood of a continuation at time t is given by A.3. Therefore, the new likelihood function for a sample of N individuals is now given by:

$$L(\beta) = \prod_{i=1}^{N} L_t^P(\beta)^{\delta_p} L_t^D(\beta)^{\delta_d} L_t^C(\beta)^{\delta_c}$$

$$= \prod_{i=1}^{N} \left\{ \frac{1}{1 + \frac{\exp(z_{di}(t)\beta_d)}{\exp(z_{pi}(t)\beta_p)}} \left(1 - \exp\left(-\left(1 + \frac{\exp(z_{di}(t)\beta_d)}{\exp(z_{pi}(t)\beta_p)}\right) \exp(z_{pi}(t)\beta_p)\right)\right) \right\}^{\delta_p}$$

$$\left\{\frac{1}{1+\frac{\exp\left(z_{pi}(t)\beta_p\right)}{\exp(z_{di}(t)\beta_d)}}\left(1-\exp\left(-\left(1+\frac{\exp\left(z_{pi}(t)\beta_p\right)}{\exp(z_{di}(t)\beta_d)}\right)\exp(z_{di}(t)\beta_d)\right)\right)\right\}^{\delta_d}$$

$$\left\{\exp\left(-\exp\left(z_{pi}(t)\beta_p\right)\right)\exp(-\exp(z_{di}(t)\beta_d))\right\}^{\delta_c}\prod_{s=1}^{t-1}\exp\left(-\exp(z_{pi}(s)\beta_p)\right)\exp(-\exp(z_{di}(s)\beta_d))$$

$$=\prod_{S=1}^{N}\prod_{i=1}^{t}\left\{\frac{1}{1+\frac{\exp(z_{di}(s)\beta_d)}{\exp\left(z_{pi}(s)\beta_p\right)}}\left(1-\exp\left(-\left(1+\frac{\exp(z_{di}(s)\beta_d)}{\exp\left(z_{pi}(s)\beta_p\right)}\right)\exp\left(z_{pi}(s)\beta_p\right)\right)\right)\right\}^{\delta_p}$$

$$\left\{\frac{1}{1+\frac{\exp\left(z_{pi}(s)\beta_p\right)}{\exp(z_{di}(s)\beta_d)}}\left(1-\exp\left(-\left(1+\frac{\exp\left(z_{pi}(s)\beta_p\right)}{\exp(z_{di}(s)\beta_d)}\right)\exp(z_{di}(s)\beta_d)\right)\right)\right\}^{\delta_d}$$

$$\left\{\exp\left(-\exp(z_{pi}(s)\beta_p)\right)\exp(-\exp(z_{di}(s)\beta_d))\right\}^{\delta_c} \qquad A.9$$

Therefore, the log-likelihood function for the entire sample is:

$$L(\beta)=\sum_{i=1}^{N}\sum_{s=1}^{t}\left\{\delta_p\left[-\ln\left(1+\frac{\exp(z_{di}(s)\beta_d)}{\exp\left(z_{pi}(s)\beta_p\right)}\right)+\ln\left(1-\exp\left(-\left(1+\frac{\exp(z_{di}(s)\beta_d)}{\exp\left(z_{pi}(s)\beta_p\right)}\right)\exp\left(z_{pi}(s)\beta_p\right)\right)\right)\right]\right.$$

$$+\delta_d\left[-\ln\left(1+\frac{\exp\left(z_{pi}(s)\beta_p\right)}{\exp(z_{di}(s)\beta_d)}\right)+\ln\left(1-\exp\left(-\left(1+\frac{\exp\left(z_{pi}(s)\beta_p\right)}{\exp(z_{di}(s)\beta_d)}\right)\exp(z_{di}(s)\beta_d)\right)\right)\right]$$

$$\left.+\delta_c\left[-\exp\left(z_{pi}(s)\beta_p\right)-\exp(z_{di}(s)\beta_d)\right]\right\}$$

$$=\sum_{i=1}^{N}\sum_{s=1}^{t}\left\{\delta_p\left[-\ln\left(1+\frac{\exp(z_{di}(s)\beta_d)}{\exp\left(z_{pi}(s)\beta_p\right)}\right)+\ln\left(1-\exp\left(-\left(\exp\left(z_{pi}(s)\beta_p\right)+\exp(z_{di}(s)\beta_d)\right)\right)\right)\right]\right.$$

$$+\delta_d\left[-\ln\left(1+\frac{\exp\left(z_{pi}(s)\beta_p\right)}{\exp(z_{di}(s)\beta_d)}\right)+\ln\left(1-\exp\left(-\left(\exp(z_{di}(s)\beta_d)+\exp\left(z_{pi}(s)\beta_p\right)\right)\right)\right)\right]$$

$$\left.+\delta_c\left[-\exp\left(z_{pi}(s)\beta_p\right)-\exp\left(z_{di}(s)\beta_d\right)\right]\right\} \qquad A.10$$

in which $z_{pi}(t), z_{di}(t)$ are vectors of time dependent explanatory variables for prepayment and default, respectively. β_p, β_d are coefficients to be estimated. Moreover, δ_p is the indicator for prepayment, δ_d is the indicator for default and δ_c is the indicator for a continuation without termination.

Appendix C

The following process is constructing the competing risks model based on the duration model introduced by Prentice and Gloeckler (1978) and Bruce D. Meyer (1990), the specification discussed by Glenn T. Sueyoshi (1992) and a class of discrete transformation survival models introduced by Adam Ding *et al* (2012). This competing risks model can be implemented in any discrete time duration analysis involving three choices in each time period; however, this study only focuses on analyzing the competing risks of prepayment and default in the single-family mortgage market.

The paper by Adam Ding *et al* (2012) showed that the survival function $S(t) = P(T \geq t), t = 1, \dots t_k$ and hence

$$\frac{S(t)}{S(t-1)} = P(T \geq t | T \geq t-1) \qquad B.1$$

and

$$G\left[-log\frac{S(t)}{S(t-1)}\right] = \exp(Z_i(t)\beta)\, G\left[-log\frac{S_0(t)}{S_0(t-1)}\right]$$

$$= \exp\big(Z_i(t)\beta + \gamma(t)\big)$$

$$= \exp(z_i(t)\beta) \qquad B.2$$

in which G was a strictly increasing transformation function with $G(0) = 0$ and $G(\infty) = \infty$, $\alpha_t = log\left[G\left[-log\frac{S_0(t)}{S_0(t-1)}\right]\right]$,$z_i(t)$is individual characteristics which contains characteristics$Z_i(t)$ and the base line, and β in the last line of B.2 contains $\gamma(t)$.

Substitute B.2 into B.1,

$$G[-\log(P(T \geq t|T \geq t-1))] = \exp(z_i(t)\beta) \qquad B.3$$

The paper by Meyer (1990) showed that the conditional survival model was:

$$P(T \geq t|T \geq t-1) = \exp(-\exp(z_i(t)\beta)) \qquad B.4$$

Substitute B.4 into B.3,

$$G[-\log(P(T \geq t|T \geq t-1))] = \exp(z_i(t)\beta)$$

$$G[-\log(\exp(-\exp(z_i(t)\beta)))] = \exp(z_i(t)\beta)$$

$$G[\exp(z_i(t)\beta)] = \exp(z_i(t)\beta)$$

$$\exp(z_i(t)\beta) = G^{-1}[\exp(z_i(t)\beta)] \qquad B.5$$

Therefore, when $G(X) = X$, $\exp(z_i(t)\beta) = \exp(z_i(t)\beta)$, and $G(X) \neq X$, substitute $\exp(z_i(t)\beta)$ by $G^{-1}[\exp(z_i(t)\beta)]$.

When there are three choices involved, continue the mortgage, prepayment and default, and the transformation function is $G_p(x)$ for prepayment and $G_d(x)$ for default. Based on B.5 and the calculation process in Appendix B, the likelihood of an event being continued at time t is

$$L_t^C(\beta) = \int_{\alpha^{*1}_t}^{\infty} \int_{\alpha^{*2}_t}^{\infty} f(u_1, u_2) du_1 du_2$$

$$= \int_{\alpha^{*}{}_t^1}^{\infty} f(u_1)du_1 \int_{\alpha^{*}{}_t^2}^{\infty} f(u_2)du_2$$

$$= \int_{\sum_{s=1}^{t} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}^{\infty} \exp(-u_1)\, du_1 \int_{\sum_{s=1}^{t} G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]}^{\infty} \exp(-u_2)\, du_2$$

$$= \exp\left(-\sum_{s=1}^{t} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]\right) \exp\left(-\sum_{s=1}^{t} G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right)$$

$$= \prod_{s=1}^{t} \exp(-G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)])\exp(-G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]) \qquad B.6$$

The likelihood of prepaying is

$$L_t^P(\beta) = \int_{\alpha^{*}{}_t^1}^{\alpha^{*}{}_t^1} \int_{g(u_1)}^{\infty} f(u_1,u_2)du_1du_2$$

$$= \int_{\sum_{s=1}^{t-1} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}^{\sum_{s=1}^{t} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]} \exp(-u_1) \int_{g(u_1)}^{\infty} \exp(-u_2)\, du_2 du_1$$

$$= \int_{\sum_{s=1}^{t-1} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}^{\sum_{s=1}^{t} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]} \exp(-u_1) \exp(-u_1 \frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}$$

$$- \sum_{s=1}^{t-1} G_d{}^{-1}[\exp(z_{di}(s)\beta_d)] + \frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]} \sum_{s=1}^{t-1} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)])du_1$$

$$= \int_{\sum_{s=1}^{t-1} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}^{\sum_{s=1}^{t} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]} \exp\left(-u_1(1 + \frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}) - \sum_{s=1}^{t-1} G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right.$$

$$\left. + \frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]} \sum_{s=1}^{t-1} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]\right) du_1$$

$$= \frac{1}{1+\frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}}\left(1-\exp\left(-\left(1+\frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}\right)G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]\right)\right)$$

$$\exp(-(1+\frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]})\sum_{s=1}^{t-1}G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]$$

$$-\sum_{s=1}^{t-1}G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]+\frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}\sum_{s=1}^{t-1}G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)])$$

$$= \frac{1}{1+\frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}}\left(1-\exp\left(-\left(1+\frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}\right)G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]\right)\right)$$

$$\exp(-\sum_{s=1}^{t-1}G_d{}^{-1}[\exp(z_{di}(s)\beta_d)])\exp(-\sum_{s=1}^{t-1}G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)])$$

$$= \frac{1}{1+\frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}}\left(1-\exp\left(-\left(1+\frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}\right)G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]\right)\right)$$

$$\prod_{s=1}^{t-1}\exp\left(-G_p{}^{-1}[\exp\left(z_{pi}(s)\beta_p\right)]\right)\exp\left(-G_d{}^{-1}[\exp\left(z_{di}(s)\beta_d\right)]\right)$$

$$= \frac{1}{1+\frac{G_d{}^{-1}[\exp\left(z_{di}(t)\beta_d\right)]}{G_p{}^{-1}[\exp\left(z_{pi}(t)\beta_p\right)]}}\left(1-\exp\left(-\left(G_p{}^{-1}[\exp\left(z_{pi}(t)\beta_p\right)]+G_d{}^{-1}[\exp\left(z_{di}(t)\beta_d\right)]\right)\right)\right)$$

$$\prod_{s=1}^{t-1}\exp\left(-G_p{}^{-1}[\exp\left(z_{pi}(s)\beta_p\right)]\right)\exp\left(-G_d{}^{-1}[\exp\left(z_{di}(s)\beta_d\right)]\right) \qquad B.7$$

in which the $g(u_1)$ function is:

$$g(u_1) = \alpha_{t-1}^2 + (u_1 - \alpha_{t-1}^1)\frac{\alpha_t^2 - \alpha_{t-1}^2}{\alpha_t^1 - \alpha_{t-1}^1}$$

$$= u_1 \frac{\alpha_t^2 - \alpha_{t-1}^2}{\alpha_t^1 - \alpha_{t-1}^1} + \alpha_{t-1}^2 - \alpha_{t-1}^1 \frac{\alpha_t^2 - \alpha_{t-1}^2}{\alpha_t^1 - \alpha_{t-1}^1}$$

$$= u_1 \frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]} + \sum_{s=1}^{t-1} G_d{}^{-1}[\exp(z_{di}(s)\beta_d)] - \sum_{s=1}^{t-1} G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)] \frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}$$

Similarly, the likelihood of default is

$$L_t^D(\beta) = \int_{\alpha^{*2}_{t-1}}^{\alpha^{*2}_t} \int_{g(u_2)}^{\infty} f(u_1, u_2) du_1 du_2$$

$$= \frac{1}{1 + \frac{G_p{}^{-1}[\exp\left(z_{pi}(t)\beta_p\right)]}{G_d{}^{-1}[\exp\left(z_{di}(t)\beta_d\right)]}} \left(1 - \exp\left(-\left(G_p{}^{-1}[\exp\left(z_{pi}(t)\beta_p\right)] + G_d{}^{-1}[\exp\left(z_{di}(t)\beta_d\right)]\right)\right)\right)$$

$$\prod_{s=1}^{t-1} \exp\left(-G_p{}^{-1}[\exp\left(z_{pi}(s)\beta_p\right)]\right) \exp\left(-G_d{}^{-1}[\exp\left(z_{di}(s)\beta_d\right)]\right) \qquad B.8$$

in which the $g(u_2)$ function is:

$$g(u_2) = \alpha_{t-1}^1 + (u_2 - \alpha_{t-1}^2)\frac{\alpha_t^1 - \alpha_{t-1}^1}{\alpha_t^2 - \alpha_{t-1}^2}$$

$$= u_2 \frac{G_p{}^{-1}[\exp\left(z_{pi}(t)\beta_p\right)]}{G_d{}^{-1}[\exp\left(z_{di}(t)\beta_d\right)]} + \sum_{s=1}^{t-1} G_p{}^{-1}[\exp\left(z_{pi}(s)\beta_p\right)] - \sum_{s=1}^{t-1} G_d{}^{-1}[\exp\left(z_{di}(s)\beta_d\right)] \frac{G_p{}^{-1}[\exp\left(z_{pi}(t)\beta_p\right)]}{G_d{}^{-1}[\exp\left(z_{di}(t)\beta_d\right)]}$$

Therefore, the likelihood function for a sample of N individuals is now given by:

$$l(\beta) = \prod_{S=1}^{N} L_t^P(\beta)^{\delta_p} L_t^D(\beta)^{\delta_d} L_t^C(\beta)^{\delta_c}$$

$$= \prod_{S=1}^{N} \left\{ \frac{1}{1 + \frac{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}} \left(1 - \exp\left(-\left(G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)] + G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]\right)\right)\right) \right\}^{\delta_p}$$

$$\left\{ \frac{1}{1 + \frac{G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)]}{G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]}} \left(1 - \exp\left(-\left(G_p{}^{-1}[\exp(z_{pi}(t)\beta_p)] + G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]\right)\right)\right) \right\}^{\delta_d}$$

$$\left\{\exp\left(- G_p{}^{-1}[\exp\left(z_{pi}(t)\beta_p\right)]\right) \exp\left(-G_d{}^{-1}[\exp(z_{di}(t)\beta_d)]\right)\right\}^{\delta_c}$$

$$\prod_{s=1}^{t-1} \exp\left(- G_p{}^{-1}[\exp\left(z_{pi}(s)\beta_p\right)]\right) \exp\left(-G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right)$$

$$= \prod_{i=1}^{N} \prod_{s=1}^{t} \left\{ \frac{1}{1 + \frac{G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}} \left(1 - \exp\left(-\left(G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)] + G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right)\right)\right) \right\}^{\delta_p}$$

$$\left\{ \frac{1}{1 + \frac{G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}{G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]}} \left(1 - \exp\left(-\left(G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)] + G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right)\right)\right) \right\}^{\delta_d}$$

$$\left\{\exp\left(- G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]\right) \exp\left(-G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right)\right\}^{\delta_c} \qquad B.9$$

in which $\delta_p$is the indicator for prepayment, $\delta_d$is the indicator for default, and $\delta_c$is the indicator for whether the mortgage continued in a certain month. Therefore, the log-likelihood function for the entire sample is

$$L(\beta) = \sum_{i=1}^{N}\sum_{s=1}^{t}\left\{\delta_p\left[-\ln\left(1+\frac{G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]}{G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}\right)\right.\right.$$

$$\left.+\ln\left(1-\exp\left(-\left(G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]+G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right)\right)\right)\right]$$

$$+\delta_d\left[-\ln\left(1+\frac{G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]}{G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]}\right)\right.$$

$$\left.+\ln\left(1-\exp\left(-\left(G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]+G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right)\right)\right)\right]$$

$$\left.+\delta_c\left[-G_p{}^{-1}[\exp(z_{pi}(s)\beta_p)]-G_d{}^{-1}[\exp(z_{di}(s)\beta_d)]\right]\right\} \qquad B.10$$

When the transformation function G is given by:

$$G_p(x) = \begin{cases}\frac{1}{c_p}\left[\exp(c_p x)-1\right], & c_p > 0\\ x, & c_p = 0\end{cases}$$

$$G_d(x) = \begin{cases}\frac{1}{c_d}\left[\exp(c_d x)-1\right], & c_d > 0\\ x, & c_d = 0\end{cases}$$

Therefore,

$$G_p{}^{-1}(x) = \begin{cases}\frac{1}{c_p}\log(1+c_p x), & c_p > 0\\ x, & c_p = 0\end{cases}$$

$$G_d{}^{-1}(x) = \begin{cases}\frac{1}{c_d}\log(1+c_d x), & c_d > 0\\ x, & c_d = 0\end{cases}$$

Hence, when $G_p(x) = x$ and $G_d(x) = x$, the log-likelihood function for the sample is the same as the function 2.12. When $\mathrm{G_p(x)} = \frac{1}{\mathrm{c_p}}\left[\exp\left(\mathrm{c_p x}\right) - 1\right]$ and $\mathrm{G_d(x)} = \frac{1}{\mathrm{c_d}}[\exp(\mathrm{c_d x}) - 1]$, the log-likelihood function for the sample is:

$$L(\beta) = \sum_{i=1}^{N}\sum_{s=1}^{t}\left\{\delta_p\left[-\ln\left(1 + \frac{\frac{1}{c_d}\log(1 + c_d\exp(z_{di}(s)\beta_d))}{\frac{1}{c_p}\log(1 + c_p\exp\left(z_{pi}(s)\beta_p\right))}\right)\right.\right.$$

$$\left. + \ln\left(1 - [1 + c_p\exp\left(z_{pi}(s)\beta_p\right)]^{-\frac{1}{c_p}}[1 + c_d\exp(z_{di}(s)\beta_d)]^{-\frac{1}{c_d}}\right)\right]$$

$$+ \delta_d\left[-\ln\left(1 + \frac{\frac{1}{c_p}\log(1 + c_p\exp\left(z_{pi}(s)\beta_p\right))}{\frac{1}{c_d}\log(1 + c_d\exp(z_{di}(s)\beta_d))}\right)\right.$$

$$\left. + \ln\left(1 - [1 + c_p\exp\left(z_{pi}(s)\beta_p\right)]^{-\frac{1}{c_p}}[1 + c_d\exp(z_{di}(s)\beta_d)]^{-\frac{1}{c_d}}\right)\right]$$

$$\left. + \delta_c\left[-\frac{1}{c_p}\log(1 + c_p\exp\left(z_{pi}(s)\beta_p\right)) - \frac{1}{c_d}\log(1 + c_d\exp(z_{di}(s)\beta_d))\right]\right\} \qquad B.11$$

Moreover, when $G_p(x) = x$ and when $G_d(x) = \frac{1}{c_d}[\exp(c_d x) - 1]$, the log-likelihood function for the sample is:

$$L(\beta) = \sum_{i=1}^{N}\sum_{s=1}^{t}\left\{\delta_p\left[-\ln\left(1+\frac{\frac{1}{c_d}\log(1+c_d\exp(z_{di}(s)\beta_d))}{\exp(z_{pi}(s)\beta_p)}\right)\right.\right.$$

$$\left.+\ln\left(1-\exp(-\exp(z_{pi}(s)\beta_p)-\frac{1}{c_d}\ln(1+c_d\exp(z_{di}(s)\beta_d)))\right)\right]$$

$$+\delta_d\left[-\ln\left(1+\frac{c_d\exp(z_{pi}(s)\beta_p)}{\log(1+c_d\exp(z_{di}(s)\beta_d))}\right)\right.$$

$$\left.+\ln\left(1-\exp(-\exp(z_{pi}(s)\beta_p)-\frac{1}{c_d}\ln(1+c_d\exp(z_{di}(s)\beta_d)))\right)\right]$$

$$\left.+\delta_c\left[-\exp(z_{pi}(s)\beta_p)-\frac{1}{c_d}\log(1+c_d\exp(z_{di}(s)\beta_d))\right]\right\} \qquad B.12$$

Appendix D

The simulation in this paper assumes the housing price stops changing after September 2004. Let CS_N indicate the new Case-Shiller index and CS_0 indicate the original Case-Shiller index. Before September 2004, the index is the original Case-Shiller index ($CS_N = CS_0$) and after September 2004, the index remains the same as the one on September 2004 ($CS_N = CS_{0\ at\ sep\ 2004}$). Therefore, the new value of the property is:

$$V_N = V_0\frac{CS_N}{CS_0} \qquad A.1$$

To keep the original loan-to-value unchanged, the loan size L_0 should change to L_N and the relationship between L_0 and L_N is

$$L_N = LTV \times V_N = LTV \times V_0\frac{CS_N}{CS_0} = L_0\frac{CS_N}{CS_0} \qquad A.2$$

Once the loan size changes, the debt-to-income ratio DTI_N also changes to $DTI_{N\,new}$ after September 2004.

$$DTI_N = \frac{L_N \frac{(\frac{r}{12}\left(1+\frac{r}{12}\right)^{360})}{(\left(1+\frac{r}{12}\right)^{360}-1)}}{I_0} = \frac{L_0 \frac{CS_N}{CS_0} \frac{(\frac{r}{12}\left(1+\frac{r}{12}\right)^{360})}{(\left(1+\frac{r}{12}\right)^{360}-1)}}{I_0} = DTI_0 \frac{CS_N}{CS_0} \qquad A.3$$

in which r is the contract rate and I_0 is the original monthly income.

Because the negative equity is calculated as the difference between the remaining balance and the property value in each month, the change of the loan size and property value leads to the change of the negative equity. The function used to calculate the new value of negative equity at month i is:

$$X_{Ni} = current\ property\ value_{Ni} - remaining\ balance_{Ni} \qquad A.4$$

$$negative\ equity_{Ni} = \begin{cases} aboslute\ value\ of\ X_{Ni} & if X_{Ni} < 0 \\ 0 & if\ X_{Ni} \geq 0 \end{cases}$$

in which, the new current property value at month i is calculated as:

$$current\ property\ value_{Ni} = \exp\left(\ln\left(\frac{CS_{Ni}}{CS_{N_0}}\right)\right) \times V_N = V_N \frac{CS_{Ni}}{CS_{N_0}} \qquad A.5$$

CS_{N_0} is the new Case-Shiller index at the note date of a mortgage and CS_{Ni} is the new Case-Shiller index at month i.

And the new remaining balance at month i is calculated as:

$$remaining\ balance_{Ni} = \frac{L_N[\left(1+\frac{r}{12}\right)^{360} - (1+\frac{r}{12})^{i}]}{\left(1+\frac{r}{12}\right)^{360} - 1} \qquad A.6$$

Reference

Joint Center for Housing Studies of Harvard University. The State of the Nation's Housing 2008. Figure 4, Page 4

Zandi Mark (2009). *Financial Shock*. FT Press. ISBN 978-0-13-701663-1

Pam Bennett (2012)."The Aftermath of the Great Recession: Financially Fragile Families and How Professionals Can Help", *The Forum for Family and Consumer Issues (FFCI)*, Vol. 17 No. 1

Jim Boswel. (2013, Jan.26th). If You Want Bank Transparency, Check Out the FDIC's Website. *Business Insider*. Retrieved from http://www.businessinsider.com/bank-transparency-via-the-fdic-2013-1

Jim Boswel. (2013, Feb.17th). Here's How Much Money US Banks Really Lost During the Crisis. *Business Insider*. Retrieved from http://www.businessinsider.com/identifying-us-banking-losses-for-the-crisis-they-helped-create-2013-2

Jim Boswel. (2013, Feb.23th). Banks' Books Still Look Ugly, But The Banking Crisis is Over. *Business Insider*. Retrieved from http://www.businessinsider.com/us-bank-loan-performance-through-the-crisis-2013-2

Mortgage-backed security from Wikipedia Retrieved from https://en.wikipedia.org/wiki/Mortgage-backed_security#cite_note-fr_stat_rel-40

Sifma Statistics. Value of Mortgage-Backed Security Issuances in $USD trillions, 1990-2009. [insert Table] Retrieved from https://en.wikipedia.org/wiki/Mortgage-backed_security#cite_note-fr_stat_rel-40

Yuliya Demyanyk, and Otto Van Hemert (2008). ”Understanding the Subprime Mortgage Crisis,” *Social Science Electronic Publishing Working Ppaers*. Retrieved 2008-10-18

Findley M.C. and D.R. Cappozza (1977). “The Variable Rate Mortgage: An Option Theory Perspective,” *Journal of Money, Credit and Banking*, 9, 356-364.

Epperson, J.F., J.B. Kau, D.C. Keenan, and W.J. Muller III (1985). “Pricing The default risk in Mortgages,” *Journal of the American Real Estate and Urban Economics Association*, 13, 261-272.

Green, J. and J.B. Shoven (1986). “The Effect of Interest Rates on Mortgage Prepayments,” *Journal of Money, Credit and Banking*, 18, 41-50.

Schwart, E.S. and W.N. Torous (1989). “Prepayment and the Valuation of Mortgage-Backed Securities,” *Journal of Finance*, 44, 375-392.

Quigley, J.M. and R. Van Order (1990). “Efficiency in the Mortgage Market: The Borrower’s Perspective,” *Journal of the American Real Estate and Urban Economics Association*, 18, 237-252.

James R. Follain, Jan Ondrich and Gyan P. Sinha (1995). “Ruthless Prepayment? Evidence from Multifamily Mortgages,” *Journal of Urban Economics*, 41, 78-101.

Quigley, J.M. and R. Van Order (1991). “Defaults on Mortgage Obligations and Capital Requirements for U.S. Savings institutions a policy perspective,” *Journal of the American Real Estate and Urban Economics Association*, 18, 237-252.

Quigley, J.M. and R. Van Order (1995). “Explicit Tests of Contingent Claims Models of Mortgage Default,” *The Journal of Real Estate Finance and Economics*, 11, 99-117

Kau, J.B., D.C. Keenan, W. J. Muller, and J. F. Epperson (1992). “A Generalized Valuation Model for Fixed-Rate Residential Mortgages,” *Journal of Money, Credit and Banking.* 24, 279-299.

--- (1995): “The Valuation at Origination of Fixed Rate Mortgages with the default and Prepayment,” *The Journal of Real Estate Finance and Economics,* 11, 5-36.

Yongheng Deng, John M. Quigley and Robert Van Order (1996). “Mortgage default and low downpayment loans: The costs of public subsidy,” *Regional Science and Urban Economics*, 23, 263-285

Yongheng Deng (1997). “Mortgage Termination: An Empirical Hazard Model with a Stochastic Term Structure,” *Journal of Real Estate Finance and Economics*. 14:3 309-331

Yongheng Deng, John M. Quigley, and Robert Van Order (2000). ”Mortgage Terminations Heterogeneity and the Exercise of Mortgage Options,” *Econometrica,* Vol.68 No.2, 275-307

Wenyi Huang and Jan Ondrich (2002). “Stay, Pay or Walk Away: A Hazard Rate Analysis of Federal Housing Administration-Insured Multifamily Mortgage Terminations,” *Journal of Housing Research*, Vol.13 No.1, 85-117

Ken Lam, Robert M. Dunsky, and Austin Kelly (2013). “Impacts of Down Payment Underwriting Standards on Loan Performance- Evidence from the GSEs and FHA Portfolios,” *FHFA Working Papers,* 13-3

Michael G. Bradley, Amy Crews Cutts and Wei Liu (2015). “Strategic Mortgage Default: The Effect of Neighborhood Factors”, *Real Estate Economics* Vol. 43 No. 2 271–299

Chao Ma (2015). "Estimating a Dynamic Discrete Choice Model with Partial Observability for Household Mortgage Default and Prepayment Behaviors," *Working Paper*

James R. Follain and Seth H. Giertz (2016). "US House Price Bubbles and Busts: Implications for Property Taxation", *Public Finance Review* Vol. 44 No. 1 132–159

William H. Greene, *Econometric Analysis*, Upper Saddle River, Prentice Hall, 2002. Print

D.R. Cox (1972). "Regression Models and Life-Tables," *Journal of the Royal Statistical Society*, Vol. 34, No. 2, pp. 187-220

D.R. Cox (1975). "Partial Likelihood," *Biometrika*, Vol. 62, pp. 269-276

R.L. Prentice and L.A. Gloeckler (1978). "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data," *Biometrics*, Vol. 34, No. 1, pp. 57-67

Bruce D. Meyer (1990). "Unemployment Insurance and Unemployment Spells," *Econometrica*, Vol. 58, No. 4, pp. 757-782

Glenn T. Sueyoshi (1992). "Semiparametric Proportional Hazards Estimation of Competing Risks Models with Time-varying Covariates," *Journal of Econometrics*, Vol. 51, pp. 25-58

A. Adam Ding, Shaonan Tian, Yan Yu and Hui Guo (2012). "A Class of Discrete Transformation Survival Models With Application to Default Probability Prediction," *Journal of the American Statistical Association*, Vol. 107 No. 499, 990-1002

Victor Aguirregabiria. (November 6, 2006). Procedure for the Maximum Likelihood estimation of McFadden's Conditional Logit, Retrieved from

http://individual.utoronto.ca/vaguirre/software/library_procedures.html

Chris Elbers and Geert Ridder (1982). “True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model,” *The Review of Economic Studies*, Vol. 49, No. 3, 403-409

J. Henckman and B. Singer (1984). “The Identifiability of the Proportional hazards Model,” *The Review of Economic Studies*, Vol. 51, 231-241

Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*, Vol II, (New York: John Wiley).

Bengt Muthen and Katherine Masyn (2005). “Discrete-Time Survival Mixture Analysis,” *Journal of Educational and Behavioral Statistics*, Vol. 30, 27-58

Ran An

Department of Economics
Maxwell School of Citizenship and Public Affairs
Syracuse University
110 Eggers Hall, Syracuse, NY 13244-1020
ranan@syr.edu | 315-560-1283

EDUCATION

Syracuse University, Maxwell School of Citizenship and Public Affairs Syracuse, NY
Doctor of Philosophy (PhD) in Economics Expected Dec. 2016

- Fields of Specialization: Econometrics, Urban Economics
- Dissertation Topics: The Competing Risks of Prepayment and Default on the Single-Family Mortgage Market

Master of Arts (MA) in Economics July 2014

- GPA: 3.49/4.00
- PhD Level Coursework: Discrete Choice and Duration Analysis, Time Series Econometrics, Econometrics I&II, Mathematics for Economists, Regional Economics, Urban Economics, Taxation, and Public Expenditure

Capital University of Economics and Business Beijing, China
Bachelor of Science (BS) in Economics July 2011

- Completed Dissertation: The Hidden Risks of the Positive Financial Policy in China
- Relevant Coursework: Finance, Econometrics, Statistics, Commercial Law, Accounting, Public Economics

CERTIFICATIONS

SAS Certified Advanced Programmer for SAS 9 Jun. 17th, 2015
SAS Certified Base Programmer for SAS 9 Mar. 9th, 2015
C Language Programmer (Grade 2 Examination) China, 2010

RESEARCH EXPERIENCE

Syracuse University Syracuse, NY
The Competing Risks on the Single-Family Mortgage Market Jan. 2014-Dec. 2016

- Construct a duration dataset based on the Single-Family mortgage dataset offered by Freddie Mac in STATA and the final sample includes more than 1 million observations.
- Collect economic indices, such as Case-Shiller indices and unemployment rates from different resources and merge them with the discrete time dataset in STATA.
- Compute important explanatory variables, such as the value of call option and the value of negative equity, in each month in GAUSS and STATA.
- Construct a competing risks model based on a multinomial logit model in STATA to analyze the competing risks of prepayment and default in Single-Family mortgage market.
- Discuss the importance of the credit score on analyzing the prepayment risk and the default risk.
- Present clear figures and tables to explain the effects of prepayment penalty, call option, unemployment rate, negative equity, original loan to value, credit score, debt-to-income ratio, loan size, loan age and house size on the probability of prepayment and the probability of default.
- Find the evidence to support the argument that outcomes for housing vary across MSAs and that the outcomes for housing are heavily influenced by local characteristics by constructing a Wald test.

New Competing Risks Models Mar. 2015- Dec. 2016

- Review the competing risks models used in the precious literature.
- Present a clear calculation process of constructing Sueyoshi competing risks model based on proportional hazard.
- Program the Sueyoshi competing risks model in GAUSS independently and the model never showed in any package.
- Argue that the coefficients estimated by the model based on Sueyoshi's proportional hazards and those estimated by the model based on the multinomial logit are insignificantly distinguishable in sign
- Present a clear calculation process of constructing a new competing risks model based on a class of discrete transformation survival models.
- Construct and program a competing risks model based on a new proportional hazard model in GAUSS program, which include the probability that one termination behavior occurs right after the other one.
- Program the new competing risks model in GAUSS independently and the model never showed in any package.
- Generalize the competing risks models by controlling different value of transformation parameters.
- Argue that the proportional hazards model is a good framework to estimate the prepayment risk but not the best framework to estimate the default and 90-days-delinquency risk.
- Argue that the coefficients estimated by the model based on the class of discrete transformation survival models are significantly different from those estimated by the other two models
- Control for unobserved heterogeneity by a latent class and construct three competing risks models.
- Program three competing risks models controlling for unobserved heterogeneity in GAUSS independently and the model never showed in any package.
- Argue that unobserved heterogeneity plays an important role in accounting for the termination risks particularly with respect to the prepayment risk.
- Construct a simulation by assuming that the Housing Price remains constant since September 2004 and argue that the dramatic increase and decrease of default and delinquency risk disappears.

WORK EXPERIENCE

Santander Bank Boston, MA
Sr. Quant Risk Analyst May 2016-Present

- Develop and support advanced regulatory-compliant credit models, including Probability of Default (PD) and Loss Given Default (LGD) and Exposure at Default (EAD) in SAS.
- Apply advanced statistical techniques to detailed credit data sourced both internally and externally.
- Work with other stakeholders both internal and external such as business and risk areas, and regulatory authorities.
- Work with the dedicated credit risk systems implementation team to support the roll-out of tactical and strategic implementations of the various credit models and methodologies.
- Contribute to model and methodology-related presentations to model committees and model users.

Syracuse University Syracuse, NY
Teaching Assistant Aug. 2011-May 2016

- Prepare Power Points and lecture 25 students in STATA program for econometric methods, a graduate level class. Lead discussions about econometric approaches used in solving different economic problems and answer questions about different economic phenomena among countries.

- Tutor students in Monetary and Banking, Financial Economics and Asset Pricing. Help students understand the foundation of monetary systems, and instruct them to solve financial problems with Microsoft Excel and C++.
- Tutor up to 200 students on Mathematical Economics, Economic Statistics, and Economic Ideas and Issues and help students improve their mathematics and statistics skills in Economics.

China Construction Bank Inner Mongolia, China
Internship Jul. 2010-Sep. 2010

- Proposed some documentation of a new fund offered by some local companies.
- Communicated with up to 5 clients every day about their investment choice, including explaining risks and return of a specific investment.

SKILLS

Computer Skills

- SQL, SAS, STATA, C language, GAUSS, Eviews, Microsoft Excel, Microsoft Power Point, Microsoft Word, Basic R language

Languages

- Chinese (Mandarin), English

HONORS

Successful Participants, Interdisciplinary Contest in Modeling United States, 2011
Honorable Mention, Interdisciplinary Contest in Modeling United States, 2010
The People's Scholarship (3 consecutive years) China, Sep. 2007-Jul. 2010
The 7th CUEB Innovation and Talent Award China, 2010
Successful Participant in China Undergraduate Mathematical Contest in Modeling China, 2009

PUBLICATIONS

Ran An. *Positive Approach of Chinese Foreign Exchange Reserve*, Chinese Urban Economy, 12, 64-66, (2010)
Ran An. *Overproduction and Modern Financial Crisis*, Securities and Futures of China, 10, 87, (2010)

PROFESSIONAL ACTIVITY AND AFFILIATIONS

Conferences
2015 Financial Management Association Annual Meeting Orlando, FL, Oct. 20th, 2015
2016 AEA Annual Meeting San Francisco, CA, Jan.3rd, 2016

Affiliations
American Economic Association

INTERESTS AND LEADERSHIP ACTIVITY

Interests

- Ballroom dance: Intermediate Social Bronze Sep.11th 2015
- Official writer on Chinese All Website 2009-Present
- Other interests: reading, traveling and singing

Leadership in Amaz Dance Group Syracuse University, 2012-2015

- Choreograph and teach traditional Chinese dance to about 20 group members
- Communicate with international student center in Syracuse University, prepared performance and documents and gave presentations to make the organization official.
- Organize the performance for Home Coming, Asian Night, etc., every semester in front of the whole university.